



956 GCAAGAGGCTACCAAGAGCTGTCGATCATTTCTG 996

seq\_name: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:US-09-126-121-22

seq\_documentation\_block:

Sequence 22, Application US/09126121

Patent No. 6252051

GENERAL INFORMATION:

APPLICANT: Godowski, Paul J., Mark, Melanle Rose, Zhang, Dong Xiao

TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/126,121

FILING DATE: 30-Jul-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Conley, Delidre L.

REGISTRATION NUMBER: 36,487

REFERENCE/DOCKET NUMBER: P1084R1D1

TELEPHONE: 650/225-2066

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 2091 base pairs

TYPE: Nucleic Acid

STRANDEDNESS: Single

TOPOLOGY: Linear

FEATURE:

NAME/KEY: Human NRG3B2 (hNRG3B2)

LOCATION: 1-2091

IDENTIFICATION METHOD:

OTHER INFORMATION:

US-09-126-121-22

US-09-877-665-4 x US-09-126-121-22

Align seg 1/1 to: US-09-126-121-22 from: 1 to: 2091

alignment\_scores:

Quality: 277.00 Length: 47

Ratio: 5.894 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-877-665-4 x US-09-126-121-22

1 HisphelyProCysArgAspLysAspLeuAlaTyrCysLeuSnaSpG1 17

856 CACTTCAAAACCTGCGGAGACAGACCTTGCATCTGCTCATGATGG 905

17 yglucysphevalileglutthrlleuthrglyserhislyshiscysargc 34

906 CGAGTCTTTGTGATCGAAACCTGACCGGATCCCATTAACACTGTCG 955

34 yslvsgluclytyrtingllyvalargcysasplnphleu 47

956 GCAAGAGGCTACCAAGAGCTGTCGATCATTTCTG 996

seq\_name: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:US-08-899-437-5

seq\_documentation\_block:

Sequence 5, Application US/08899437

Patent No. 6121415

GENERAL INFORMATION:

APPLICANT: Godowski, Paul J., Mark, Melanle Rose, Zhang, Dong Xiao

TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/899,437

FILING DATE: 24-Jul-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Conley, Delidre L.

REGISTRATION NUMBER: 36,487

REFERENCE/DOCKET NUMBER: P1084R1

TELEPHONE: 650/225-2066

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 2502 base pairs

TYPE: Nucleic Acid

STRANDEDNESS: Single

TOPOLOGY: Linear

FEATURE:

NAME/KEY: Human NRG3B1(hNRG3B1)/nucleic acid seq.

LOCATION: 1-2502

IDENTIFICATION METHOD:

OTHER INFORMATION:

US-08-899-437-5

US-08-899-437-5

Align seg 1/1 to: US-08-899-437-5 from: 1 to: 2502

alignment\_scores:

Quality: 277.00 Length: 47

Ratio: 5.894 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-877-665-4 x US-08-899-437-5

1 HisphelyProCysArgAspLysAspLeuAlaTyrCysLeuSnaSpG1 17

999 CACTTCAAAACCTGCGGAGACAGACCTTGCATCTGCTCATGATGG 1048

17 yglucysphevalileglutthrlleuthrglyserhislyshiscysargc 34

1049 CGAGTCTTTGTGATCGAAACCTGACCGGATCCCATTAACACTGTCG 1098

34 yslvsgluclytyrtingllyvalargcysasplnphleu 47

1099 GCAAGAGGCTACCAAGAGCTGTCGATCATTTCTG 1139

seq\_name: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:US-09-126-121-5

seq\_documentation\_block:

Sequence 5, Application US/09126121

Patent No. 6252051

GENERAL INFORMATION:

APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
 TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
 NUMBER OF SEQUENCES: 23  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Genentech, Inc.  
 STREET: 1 DNA Way  
 CITY: South San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94080  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Winpatin (Genentech)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/126,121  
 FILING DATE: 30-Jul-1998  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Conley, Delidre L.  
 REGISTRATION NUMBER: 36,487  
 REFERENCE/DOCKET NUMBER: P1084R1D1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 650/225-2066  
 TELEFAX: 650/952-9881  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2502 base pairs  
 TYPE: Nucleic Acid  
 STRANDEDNESS: Single  
 TOPOLOGY: Linear  
 FEATURE:  
 NAME/KEY: Human NR3B1(hNR3B1)/nucleic acid seq.  
 LOCATION: 1-2502  
 IDENTIFICATION METHOD:  
 OTHER INFORMATION:  
 US-09-126-121-5  
 alignment\_scores:  
 Quality: 277.00 Length: 47  
 Ratio: 5.894 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000  
 alignment\_block:  
 US-09-877-665-4 x US-09-126-121-5 ..  
 Align seg 1/1 to: US-09-126-121-5 from: 1 to: 2502  
 1 H1SPHELYPPOCYARGAPLYSAPLEUALATYRCYSLAUNASPC1 17  
 |||||||  
 999 CACTTCAAAACCTGCGAGACAAGACCTTGCTACTGCTCAATGATG 1048  
 17 YGLUCYSHEVALIIIEGLUTHLEUTHGLYSERHISYSHSCYSARGC 34  
 |||||||  
 1049 CGAGGCTTTGGATGCAAGACCTGACCGGATCCATAAACAATCGGT 1098  
 34 YSLVSLGLUGLYTYRGLINGLYVALARGCYASPGINPHELEU 47  
 |||||||  
 1099 GCAGAGAGAGGCTACCAAGAGATCCCTGTGATCAATTTCTG 1139  
 seq\_name: /cgn2\_6/ptodata/1/1na/6A\_COMB.seq:US-08-899-437-1  
 seq\_documentation\_block:  
 Sequence 1, Application US/08899437  
 Patent No. 6121415  
 GENERAL INFORMATION:  
 APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
 TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
 NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Genentech, Inc.  
 STREET: 1 DNA Way  
 CITY: South San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94080  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Winpatin (Genentech)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/899,437  
 FILING DATE: 24-Jul-1997  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Conley, Delidre L.  
 REGISTRATION NUMBER: 36,487  
 REFERENCE/DOCKET NUMBER: P1084R1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 650/225-2066  
 TELEFAX: 650/952-9881  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2538 base pairs  
 TYPE: Nucleic Acid  
 STRANDEDNESS: Single  
 TOPOLOGY: Linear  
 FEATURE:  
 NAME/KEY: mouse NR3 nucleic acid  
 LOCATION: 1-2538  
 IDENTIFICATION METHOD:  
 OTHER INFORMATION:  
 US-08-899-437-1  
 alignment\_scores:  
 Quality: 277.00 Length: 47  
 Ratio: 5.894 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000  
 alignment\_block:  
 US-09-877-665-4 x US-08-899-437-1 ..  
 Align seg 1/1 to: US-08-899-437-1 from: 1 to: 2538  
 1 H1SPHELYPPOCYARGAPLYSAPLEUALATYRCYSLAUNASPC1 17  
 |||||||  
 1150 CACTTCAAAACCTGCGAGACAAGACCTGCGTATGCTCAATGATG 1199  
 17 YGLUCYSHEVALIIIEGLUTHLEUTHGLYSERHISYSHSCYSARGC 34  
 |||||||  
 1200 TGAATGCTTTGGATGAGACCTGACAGAGATCCATAAACAATCGGT 1249  
 34 YSLVSLGLUGLYTYRGLINGLYVALARGCYASPGINPHELEU 47  
 |||||||  
 1250 GCAGAGAGAGGCTACCAAGAGATCCCTGTGATCAATTTCTG 1290  
 seq\_name: /cgn2\_6/ptodata/1/1na/6B\_COMB.seq:US-09-126-121-1  
 seq\_documentation\_block:  
 Sequence 1, Application US/09126121  
 Patent No. 6252051  
 GENERAL INFORMATION:  
 APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
 TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
 NUMBER OF SEQUENCES: 23  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Genentech, Inc.  
 STREET: 1 DNA Way  
 CITY: South San Francisco

```

STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/126,121
FILING DATE: 30-JUL-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Conley, Delidre L.
REGISTRATION NUMBER: 36,487
REFERENCE/DOCKET NUMBER: P1084R1D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-2066
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2538 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
FEATURE:
NAME/KEY: mouse NRG3 nucleic acid
LOCATION: 1-2538
IDENTIFICATION METHOD:
OTHER INFORMATION:
US-09-126-121-1

```

```

alignment_scores:
Quality: 277.00      Length: 47
Ratio: 5.894        Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000

```

```

alignment_block:
US-09-877-665-4 x US-09-126-121-1

```

```

Align seg 1/1 to: US-09-126-121-1 from: 1 to: 2538

```

```

1 HisphelyProCysArGAspLysAspLeuAlaTyrCysLeuAsnAspG1 17
|||||
1150 CACTTCAACCTGTGAGACAGAGACCTGGCGTATGTCTCAATGATGG 1199
17 yGIuCySpheValIleGIuThrLeuThrGIySerHisLysHisCysArgC 34
|||||
1200 TGAATGCTTGTGATTGAGACCTGACAGGATCCATTAAGCAGCTCTCGT 1249
34 yLySGlUGlYTyRGInGlyValAlaGcYsAspGlnPhelEu 47
|||||
1250 GGAAGGAAGGCTACCAAGAGAGCCGTGTGATCAATTTCTG 1290

```

```

seq_name: /cgn2_6/ptodata/1/lna/5A_COMB.seq:US-09-238-182-2

```

```

seq_documentation_block:
; Sequence 2, Application US/09238182
; Patent No. 6080845
; GENERAL INFORMATION:
; APPLICANT: Carnahan, Josette F.
; TITLE OF INVENTION: MONOCLONAL ANTIBODY AGAINST UTRICULAR EPITHELIUM
; FILE REFERENCE: A-532A
; CURRENT APPLICATION NUMBER: US/09/238,182
; EARLIER FILING DATE: 1999-01-28
; EARLIER APPLICATION NUMBER: 09/129,549
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 156
; TYPE: DNA

```

```

; ORGANISM: Human
US-09-238-182-2

```

```

alignment_scores:
Quality: 116.50      Length: 48
Ratio: 3.149        Gaps: 1
Percent Similarity: 77.083  Percent Identity: 33.333

```

```

alignment_block:
US-09-877-665-4 x US-09-238-182-2

```

```

Align seg 1/1 to: US-09-238-182-2 from: 1 to: 156

```

```

1 HisphelyProCysArGAspLysAspLeuAlaTyrCysLeuAsnAspG1 17
|||||
4 CACTTGTAAATGTGCGGAGAGAGAAACTTCTGTGTGATGAGG 53
17 yGIuCySpheValIleGIuThrLeuThrGIySerHisLysHis...CysA 33
|||||
54 GAGAGCTCTCATGTGAGAGACCTTCAACCCCTGAGATCTGTGCA 103
33 yGcYsLysGluGlyTyRGInGlyValAlaGcYsAspGlnPhelEu 47
|||||
104 AGTCCACCTGATCTACTGAGAGCAAGATGCAAACTACGTA 147

```

```

seq_name: /cgn2_6/ptodata/1/lna/5A_COMB.seq:US-08-036-555B-155

```

```

seq_documentation_block:
; Sequence 155, Application US/08036555B
; Patent No. 5530109
; GENERAL INFORMATION:

```

```

; APPLICANT: Goodheart, Andrew; Stroobant, Paul;
; APPLICANT: Minchetti, Luisa; Waterfield, Michael; Marchioni, Mark;
; APPLICANT: Chen, Mao Su; Hiles, Ian
; TITLE OF INVENTION: Glial Mitogenic Factors, Their
; NUMBER OF SEQUENCES: 184

```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Feltz & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

```

```

; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WordPerfect

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/036,555B
; FILING DATE: 24-MAR-1993

```

```

; CLASSIFICATION: 435

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/965,173

```

```

; FILING DATE: 23-OCT-1992

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/940,389

```

```

; FILING DATE: 03-SEP-1992

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/907,138

```

```

; FILING DATE: 30-JUN-1992

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/863,703

```

```

; FILING DATE: 03-APRIL-1992

```

```

; APPLICATION NUMBER: U.K. 91 07566.3

```

```

; FILING DATE: 10-APRIL-1991

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Tsai, Christine H.

```

```

; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: LDD 5250.4

```

```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 155:
SEQUENCE CHARACTERISTICS:
LENGTH: 192
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-036-555B-155

alignment_scores:
Quality: 113.50      Length: 46
Ratio: 3.338         Gaps: 1
Percent Similarity: 73.913      Percent Identity: 34.783

alignment_block:
US-09-877-665-4 x US-08-036-555B-155

Align seg 1/1 to: US-08-036-555B-155 from: 1 to: 192

1 HspHelysProCysArgAspLysAspLeuAlaTyrCysLeuAsnAspG1 17
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
4 CATCTGTGCAAGTGTGCAGAGAGAGAAACTTCTGTGTGAATGAGAG 53
17 yGlucysPheValIleGluThrLeuThrGlySerHisHis...CysA 33
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
54 CGAGTGTCTCATGTGTGAAGAGACCTTCAATCCCTCAAGAACTGTGCA 103
33 rGcYsLysGluGlyTyrGlnGlyValArgCysAspGln 45
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
104 AGTGCACACCTGATGATGAGAGAGAGATGACTGAG 141

seq_name: /cgn2_6/ptodata/1/lna/5A_COMB.seq:US-08-469-569-155

seq_documentation_block:
Sequence 155, Application US/08469569
Patent No. 5606032
GENERAL INFORMATION:
APPLICANT: Goodheart, Andrew; Stroobant, Paul;
APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchionl, Mark;
APPLICANT: Chen, Malo Su; Hiles, Ian
TITLE OF INVENTION: Glial Mitogenic Factors, Their
TITLE OF INVENTION: Preparation and Use
NUMBER OF SEQUENCES: 184
CORRESPONDENCE ADDRESS:
ADDRESS: Felie & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 Inch, 360 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,569
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/036,555
FILING DATE: 24-MAR-1993
APPLICATION NUMBER: 07/965,173
FILING DATE: 23-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/940,389
FILING DATE: 03-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/907,138
FILING DATE: 30-JUN-1992

```

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/863,703
FILING DATE: 03-APRIL-1992
PRIOR APPLICATION DATA: U.K. 91 07566.3
APPLICATION NUMBER: 10-APRIL-1991
FILING DATE: 10-APRIL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Tsai, Christine H.
REGISTRATION NUMBER: 34,266
REFERENCE/DOCKET NUMBER: LUD 5250.4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 155:
SEQUENCE CHARACTERISTICS:
LENGTH: 192
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-469-569-155

alignment_scores:
Quality: 113.50      Length: 46
Ratio: 3.338         Gaps: 1
Percent Similarity: 73.913      Percent Identity: 34.783

alignment_block:
US-09-877-665-4 x US-08-469-569-155

Align seg 1/1 to: US-08-469-569-155 from: 1 to: 192

1 HspHelysProCysArgAspLysAspLeuAlaTyrCysLeuAsnAspG1 17
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
4 CATCTGTGCAAGTGTGCAGAGAGAGAAACTTCTGTGTGAATGAGAG 53
17 yGlucysPheValIleGluThrLeuThrGlySerHisHis...CysA 33
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
54 CGAGTGTCTCATGTGTGAAGAGACCTTCAATCCCTCAAGAACTGTGCA 103
33 rGcYsLysGluGlyTyrGlnGlyValArgCysAspGln 45
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
104 AGTGCACACCTGATGATGAGAGAGAGATGACTGAG 141

seq_name: /cgn2_6/ptodata/1/lna/5A_COMB.seq:US-08-249-322A-155

seq_documentation_block:
Sequence 155, Application US/08249322A
Patent No. 5716930
GENERAL INFORMATION:
APPLICANT: Goodheart, Andrew; Stroobant, Paul;
APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchionl, Mark;
APPLICANT: Chen, Malo Su; Hiles, Ian
TITLE OF INVENTION: Glial Mitogenic Factors, Their
TITLE OF INVENTION: Preparation and Use
NUMBER OF SEQUENCES: 184
CORRESPONDENCE ADDRESS:
ADDRESS: Felie & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 Inch, 360 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/249,322A
FILING DATE: 26-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

```

Align seg 1/1 to: US-08-249-322A-155 from: 1 to: 192

```
seq_name: /cgn2_0/plodata/1/lna/5A_COMB.seq:US-08-469-526A-155
seq_documentation_block:
: Sequence information
```

GENERAL INFORMATION:  
APPLICANT: Goodheart, Andrew  
APPLICANT: Strobant, Paul  
APPLICANT: Mingnelli, Luisa  
APPLICANT: Waterfield, Michel  
APPLICANT: Marchionni, Mark  
APPLICANT: Chen, Maio Su  
APPLICANT: Hiles, Ian  
TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR  
TITLE OF INVENTION: PREPARATION AND USE  
NUMBER OF SEQUENCES: 187  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP

```

alignment_scores:      Quality: 113.50      Length: 46
                       Ratio: 3.338      Gaps: 1
Percent Similarity:    73.813      Percent Identity: 34.793

alignment block:
US-09-877-665-4 x US-08-469-526A-155  ..

```

Align seg 1/1 to: US-08-469-526A-155 from: 1 to: 192

```

1 HisPheLysProCysArgAspLysAspLeuAlaTyrCysLeuAsnAspG1 17
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4 CATCTTCATCAAGTGGCAGAAAGAGAGAAAACCTTCTGTGTGAATGGAG 53
17 yGluCysPheValIleGluThrLeuThrGlySerHisLysHis...CysA 33
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
54 CGAATGCTTCAATGGTGAAGACCTTTCAAAATCCCTCAAGATACTTGTGCA 103
33 rGcysLysGluGlyTyrGluGlnValAlaArgCysAspGln 45
   ::|||:: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
104 AGTGCACAACCTGGATTACTGTGAGCCGACGAGATGTACTGTGAG 141

```

seq\_name: /cgn2\_6/ptodata/1/lna/5B\_COMB.seq:US-08-734-591A-155

```
seq documentation block:
: Sequence 155. Application US/08734591A
: Patent No. 3854220
:
: GENERAL INFORMATION:
:
: APPLICANT: Goodearl, Andrew
: APPLICANT: Stroobant, Paul
```

APPLICANT: Minghetti, Luisa  
APPLICANT: Waterfield, Michael  
APPLICANT: Hiles, Ian  
APPLICANT: Marchionni, Mark  
APPLICANT: Chen, Mario  
TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR  
NUMBER OF SEQUENCES: 187  
PREPARATION AND USE  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbling LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM Compatible Pentium  
OPERATING SYSTEM: Windows95  
SOFTWARE: WordPerfect (Version 7.0)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/734,591A  
FILING DATE: 22-OCT-1996  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/470,335  
FILING DATE: 06-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/036,555  
FILING DATE: 03-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/965,173  
FILING DATE: 23-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/940,389  
FILING DATE: 03-SEP-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/907,138  
FILING DATE: 30-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/863,703  
FILING DATE: 03-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: UK 91 07566.3  
FILING DATE: 10-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Bleker-Brady, Kristina  
REGISTRATION NUMBER: 39,109  
REFERENCE/DOCKET NUMBER: 04585/00200P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 428-0200  
TELEFAX: (617) 428-7045  
TELEX:  
INFORMATION FOR SEQ ID NO: 155:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-734-591A-155  
alignment\_scores:  
Quality: 113.50 Length: 46  
Ratio: 3.338 Gaps: 1  
Percent Similarity: 73.913 Percent Identity: 34.783  
alignment\_block:  
US-09-877-665-4 x US-08-734-591A-155 ..  
Align seg 1/1 to: US-08-734-591A-155 from: 1 to: 192  
1 HAsphelysProcysarGAspLysAspLeuAlaTyrCysLeuAsnAspG1 17

|||||  
4 CACCTGTCAGAGTGCAGAGAGAGAAAACTTCTGTGTCATGAG 53  
17 ygluGspheValIleGlutThrleuThrglySerHis...CysA 33  
|||||  
54 CGAGTCTTCATGTCGTAAGACCTTCAATCCCTCAAGATACTTGTGCA 103  
33 rGcYslsglGlyTyrGlnGlyValArgCysAspGln 45  
|||||  
104 AGTCCACCTGATGATCTACTGAGAGAGATGATCTGAG 141

seq\_name: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:US-08-469-660-155

seq\_documentation\_block:  
; Sequence 155, Application US/08469660  
; Patent No. 5876973  
; GENERAL INFORMATION:  
; APPLICANT: Gwynne, David I.; Marchionni, Mark;  
; APPLICANT: McBurney, Robert N.  
; TITLE OF INVENTION: INHIBITORS OF CELL PROLIFERATION,  
; TITLE OF INVENTION: THEIR PREPARATION AND USE  
; NUMBER OF SEQUENCES: 184  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; ZIP: 0211-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
; COMPUTER: IBM  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: WordPerfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/469,660  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/011,396  
; FILING DATE: 29-JAN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/984,085  
; FILING DATE: 01-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/951,747  
; FILING DATE: 25-SEP-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/927,337  
; FILING DATE: 10-AUG-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 04585/017004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: 200154  
; INFORMATION FOR SEQ ID NO: 155:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 192  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; US-08-469-660-155  
alignment\_scores:  
Quality: 113.50 Length: 46  
Ratio: 3.338 Gaps: 1  
Percent Similarity: 73.913 Percent Identity: 34.783  
alignment\_block:  
US-09-877-665-4 x US-08-469-660-155 ..

Align seg 1/1 to: US-08-469-660-155 from: 1 to: 192

```

1 HisPheLysProCysArgAspLysAspLeuAlaTyrcysLeuAsnAspG1 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4 CATCTGTCAAGTGTGCAGAGAGAGAGAAACTTCTGTGTGATGAGG 53
17 yglucysPheValIleGluThrLeuThrGlySerHisLysHis...CysA 33
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
54 CGAGTCTTCATGTGTGAAGACCTTCAATCCCTCAAGATCTGTGCA 103
33 rgcysLysgluGlyTyrcinglyValArgCysAspGln 45
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
104 AGTCCCAACCTGGATTCTACTGAGCGAGCATGTACTGAG 141

```

seq\_name: /cgn2\_6/ptodata/1/lna/6A\_COMB.seq:US-08-341-018-61

seq\_documentation\_block:

```

; Sequence 61, Application US/08341018A
; Patent No. 6087323
; GENERAL INFORMATION:
; APPLICANT: Gwynne, David I.
; APPLICANT: Marchionni, Mark A.
; APPLICANT: Mahanthappa, Nagesh K.
; APPLICANT: Birmingham, Mark A.
; APPLICANT: McBurney, Robert N.
; APPLICANT: Goldin, Stanley M.
; TITLE OF INVENTION: USE OF NEUREGUIN AS MODULATORS OF
; FILE REFERENCE: 04585/041001
; CURRENT APPLICATION NUMBER: US/08/341,018A
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 192
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(189)
US-08-341-018-61

```

alignment\_scores:

```

Quality: 113.50 Length: 46
Ratio: 3.338 Gaps: 1
Percent Similarity: 73.913 Percent Identity: 34.783

```

alignment\_block:

US-09-877-665-4 x US-08-341-018-61 ..

Align seg 1/1 to: US-08-341-018-61 from: 1 to: 192

```

1 HisPheLysProCysArgAspLysAspLeuAlaTyrcysLeuAsnAspG1 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4 CATCTGTCAAGTGTGCAGAGAGAGAGAAACTTCTGTGTGATGAGG 53
17 yglucysPheValIleGluThrLeuThrGlySerHisLysHis...CysA 33
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
54 CGAGTCTTCATGTGTGAAGACCTTCAATCCCTCAAGATCTGTGCA 103
33 rgcysLysgluGlyTyrcinglyValArgCysAspGln 45
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
104 AGTCCCAACCTGGATTCTACTGAGCGAGCATGTACTGAG 141

```

seq\_name: /cgn2\_6/ptodata/1/lna/6A\_COMB.seq:US-08-470-335-155

seq\_documentation\_block:

```

; Sequence 155, Application US/08470335F
; Patent No. 6147190
; GENERAL INFORMATION:
; APPLICANT: GOODEARL, ANDREW
; APPLICANT: STROOBANT, PAUL

```

```

; APPLICANT: MINCHETTI, LUISA
; APPLICANT: WATERFIELD, MICHAEL
; APPLICANT: MARCHIONNI, MARK
; APPLICANT: CHEN, MARIO S.
; APPLICANT: HILES, IAN

```

```

; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
; FILE REFERENCE: 04585/00200B
; CURRENT APPLICATION NUMBER: US/08/470,335F
; CURRENT FILING DATE: 1995-06-06
; EX. IER APPLICATION NUMBER: 08/036,555
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 155
; LENGTH: 192
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(189)
US-08-470-335-155

```

alignment\_scores:

```

Quality: 113.50 Length: 46
Ratio: 3.338 Gaps: 1
Percent Similarity: 73.913 Percent Identity: 34.783

```

alignment\_block:

US-09-877-665-4 x US-08-470-335-155 ..

Align seg 1/1 to: US-08-470-335-155 from: 1 to: 192

```

1 HisPheLysProCysArgAspLysAspLeuAlaTyrcysLeuAsnAspG1 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4 CATCTGTCAAGTGTGCAGAGAGAGAGAAACTTCTGTGTGATGAGG 53
17 yglucysPheValIleGluThrLeuThrGlySerHisLysHis...CysA 33
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
54 CGAGTCTTCATGTGTGAAGACCTTCAATCCCTCAAGATCTGTGCA 103
33 rgcysLysgluGlyTyrcinglyValArgCysAspGln 45
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
104 AGTCCCAACCTGGATTCTACTGAGCGAGCATGTACTGAG 141

```







CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple  
 CC sclerosis, Huntington's chorea, Down's syndrome, nerve deafness,  
 CC and Meniere's disease. They can also be used to treat neuropathies  
 CC associated with systemic disease including post-polio syndrome,  
 CC hereditary neuropathies including Charcot-Marie-Tooth disease,  
 CC Retsum's disease, abetalipoproteinemia, Tangier disease, Krabbe's  
 CC disease, metachromatic leukodystrophy, Fabry's disease and  
 CC Dejerine-Sottas syndrome, to treat disease of skeletal muscle of  
 CC smooth muscle, such as muscular dystrophy or diseases caused by  
 CC skeletal or smooth muscle wasting. The products can also be used  
 CC for detection, diagnosis, for the production of transgenic or  
 CC knockout animals or for drug screening.

XX Sequence 2091 BP; 528 A; 633 C; 515 G; 415 T; 0 other;

#### Alignment\_scores:

Quality:	277.00	Length:	47
Ratio:	5.894	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

#### Alignment\_block:

US-09-877-665-4 x AAX06989

Align seg 1/1 to: AAX06989 from: 1 to: 2091

```

1 HisphelyProCySArGAspLySApLeuAlaTYrCYSLeA-nAspG1 17
856 CACTTCAAAACCTGCGGAGACAGACCTTGCTGCTCATGATGATG 905
17 YGlucSpheValIleGluThrLeuThrGlySerHisLysHisCYsArG 34
906 CGAGTGCCTTTGTGATCGAAACCTGACCGGATCCATTAACACGTGCGT 955
34 YSLySGlUGlTYrGlnGlyValAlrGcYsAspGlnPheLeu 47
956 GCAAGAAGAGGCTACCAAGAGTCCGTTGTGATCAATTTCTG 996

```

seq\_name: /SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA1999.DAT: AAX36423

#### seq\_documentation\_block:

ID AAX36423 standard; DNA; 2199 BP.

XX AAX36423;

DT 06-JUL-1999 (first entry)

DE Human heregulin-like factor coding sequence.

KW Human heregulin-like factor; HLF; cell growth regulator; diagnosis;  
 KW neural system disorder; cancer; ss.

OS Homo sapiens.

PN WO9857989-A1.

PD 23-DEC-1998.

PF 16-JUN-1998; 98WO-US12403.

PR 17-JUN-1997; 97US-0049942.

PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (GEOU) UNIV GEORGETOWN.

PI H1j4z1 MM, King CR, Ruben SM, Young P;

DR WPI; 1999-095327/08.

DR P-PSDB; AAY05451.

PT New isolated heregulin-like factor - used to develop products for  
 PT the diagnosis and treatment of disorders involving regulation of  
 PT cell growth, particularly cancers

XX Claim 2; Page 86-87; 118pp; English.

PS This sequence encodes the human heregulin-like factor (HLF) of the  
 XX invention. The HLF is involved in the regulation of cell growth.  
 CC Detection of different levels of expression of the HLF gene can be used  
 CC for the diagnosis of disorders, e.g. in the neural system. In  
 CC particular, detection of different levels of HLF gene expression in cells  
 CC or body fluid of an individual can be used for diagnosing cancer. The  
 CC products can also be used in the treatment of disorders involving  
 CC abnormal levels of HLF activity.

XX Sequence 2199 BP; 689 A; 445 C; 462 G; 603 T; 0 other;

#### Alignment\_scores:

Quality:	277.00	Length:	47
Ratio:	5.894	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

#### Alignment\_block:

US-09-877-665-4 x AAX36423

Align seg 1/1 to: AAX36423 from: 1 to: 2199

```

1 HisphelyProCySArGAspLySApLeuAlaTYrCYSLeA-nAspG1 17
92 CACTTCAAAACCTGCGGAGACAGACCTTGCTGCTCATGATGATG 141
17 YGlucSpheValIleGluThrLeuThrGlySerHisLysHisCYsArG 34
142 CGAGTGCCTTTGTGATCGAAACCTGACCGGATCCATTAACACGTGCGT 191
34 YSLySGlUGlTYrGlnGlyValAlrGcYsAspGlnPheLeu 47
192 GCAAGAAGAGGCTACCAAGAGTCCGTTGTGATCAATTTCTG 232

```

seq\_name: /SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA1999.DAT: AAX06988

#### seq\_documentation\_block:

ID AAX06988 standard; CDNA; 2502 BP.

XX AAX06988;

DT 10-MAY-1999 (first entry)

DE Human heregulin related ligand NRG3 CDNA.

KW Neuregulin related ligand; NRG3; hNRG3B1; human; ErbB4 receptor;  
 KW signal transduction; nervous system disorder; neurodegeneration;  
 KW neuropathy; therapy; diagnosis; ss.

OS Homo sapiens.

PN WO9902681-A1.

PD 21-JAN-1999.

PF 30-JUN-1998; 98WO-US13411.

PR 24-JUL-1997; 97US-0899437.

PR 09-JUL-1997; 97US-0052019.

PA (GETH) GENENTECH INC.

PI Godowski PJ, Mark MR, Zhang D;

DR WPI; 1999-120882/10.

DR P-PSDB; AAW97618.

XX New isolated neuroguilin related ligand-3 - used to develop products  
 PT for treating nervous system disorders, e.g. stroke, ischaemia,  
 PT infection, malignancy, Alzheimer's disease or Down's syndrome  
 XX  
 PS Example 1; Page 64-66; 101pp: English.

CC This cDNA clone encodes human neuroguilin related ligand NR3 (see  
 CC AAW97618), a novel member of the epidermal growth factor (EGF)-like  
 CC family of protein ligands that binds to the ErbB4 receptor, but not  
 CC to the ErbB2 or ErbB3 receptor, and which activates ErbB4 receptor  
 CC tyrosine phosphorylation. A partial human NR3 cDNA clone was  
 CC obtained from a foetal brain cDNA library using a probe (see  
 CC AAX06987) based on an isolated expressed sequence tag sequence (see  
 CC AAX06990). Additional 5' sequence was obtained by anchored PCR. A  
 CC nucleic acid comprising the human NR3 open reading frame cloned in  
 CC an expression vector is designated PRK5.tk.neo.NNRG3B1 (ATCC 209157).  
 CC An alternatively spliced form of human NR3 nucleic acid, hNRG3B2,  
 CC lacking nucleotides 1585-1656, is given in AAX06989. The chromosomal  
 CC localisation of human NR3 was mapped to 10q22. The invention  
 CC provides human and murine NR3 polynucleotides (see also AAX06987) and  
 CC polypeptides (see also AAW97617), expression vectors, host cells and  
 CC methods for the recombinant production of NR3s. The polynucleotides  
 CC and polypeptides can be used to enhance the survival, proliferation  
 CC or differentiation of cells having the ErbB4 receptor in vivo and in  
 CC vitro. They can be used to prevent or treat damage to a nerve or  
 CC damage to other NR3-expressing or NR3-responsive cells, e.g.  
 CC brain, heart, or kidney cells. In particular, they can be used to  
 CC treat diseases which involve neural cell growth such as demyelination,  
 CC or damage or loss of glial cells (e.g. multiple sclerosis). They can  
 CC be used to treat patients whose nervous system has been damaged by  
 CC e.g. trauma, surgery, stroke, ischaemia, infection, metabolic  
 CC disease, nutritional deficiency, malignancy, or toxic agents. NR3  
 CC can also be used to treat motor neuron disorders such as amyotrophic  
 CC lateral sclerosis (Lou Gehrig's disease), Bell's palsy, and  
 CC conditions involving spinal muscular atrophy or paralysis. NR3  
 CC can be used to treat human neurodegenerative disorders such as  
 CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple  
 CC sclerosis, Huntington's chorea, Down's syndrome, nerve deafness,  
 CC and Meniere's disease. They can also be used to treat neuropathies  
 CC associated with systemic disease including post-polio syndrome,  
 CC hereditary neuropathies including Charcot-Marie-Tooth disease, Krabbe's  
 CC Refsum's disease, abetalipoproteinemia, Fabry's disease and  
 CC disease, metachromatic leukodystrophy, Fanger's disease and  
 CC Dejerine-Sottas syndrome, to treat disease or diseases caused by  
 CC skeletal or smooth muscle wasting. The products can also be used  
 CC for detection, diagnosis, for the production of transgenic or  
 CC knockout animals or for drug screening.

XX Sequence 2502 BP; 647 A; 723 C; 614 G; 517 T; 1 other;

alignment\_scores:

Quality: 277.00 Length: 47  
 Ratio: 5.894 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-877-665-4 x AAX06988

Align seg 1/1 to: AAX06988 from: 1 to: 2502

1 H1SPHELSPROCSATGASPLYSASPLAULATYRCYSLEUAANAAPG1 17  
 999 CACTTCAAAACCTCCGACGACAGACCTTGCACATCTCTCAATGATG 1048  
 17 YGLUCSPHEVAL1LEGLUTHRTLEUTHRTGLYSERH1SLYSH1SCYSARGC 34  
 1049 CGAGTGTCTTGTGATCGAAACCCCTGACCCGATCCCATTAACACATGTCGGT 1098  
 34 YSLYSGLUGLTYTGTGLNGLYVALARGCYSAPG1NPHLEU 47

1099 GCAAAGAGCGCTACCAAGAGATCCGTTGTGATCAATTCTG 1139

seq\_name: /SID51/gcgdata/geneseq/geneseq-emb1/NA1999.DAT: AAX06987

seq\_documentation\_block:

ID AAX06987 standard; CDNA: 2538 BP.

AC AAX06987;

DT 10-MAY-1999 (first entry)

DE Mouse neuroguilin related ligand NR3 cDNA.

KW Neuroguilin related ligand; NR3; mouse; ErbB4 receptor;

KW signal transduction; nervous system disorder; neurodegeneration;

KW neuropathy; therapy; diagnosis; ss.

OS Mus sp.

FT Key Location/Qualifiers

FT CDS 289..2430

FT /tag- a

PN R09902681-A1.

PF 30-JUN-1998; 98MO-US13411.

PR 24-JUL-1997; 97US-0899437.

PR 09-JUL-1997; 97US-0052019.

PA (GETH ) GENENTECH INC.

PI Godowski PJ, Mark MR, Zhang D;

DR NPI; 1999-120882/10.

DR P-PDB; AAW97617.

PT New isolated neuroguilin related ligand-3 - used to develop products

PT for treating nervous system disorders, e.g. stroke, ischaemia,

PT infection, malignancy, Alzheimer's disease or Down's syndrome

XX Example 1; Page 57-59; 101pp: English.

XX This cDNA clone encodes murine neuroguilin related ligand NR3 (see  
 CC AAW97617), a novel member of the epidermal growth factor (EGF)-like  
 CC family of protein ligands that binds to the ErbB4 receptor, but not  
 CC to the ErbB2 or ErbB3 receptor, and which activates ErbB4 receptor  
 CC tyrosine phosphorylation. To clone murine NR3 cDNA, degenerate  
 CC primers based on regions (see AAW97623-24) proximal to the  
 CC transmembrane domain of a partial human cDNA were used to screen  
 CC a mouse brain cDNA library. An isolated clone was used as a  
 CC probe to obtain 2 further, overlapping clones that provided the  
 CC entire open reading frame. A nucleic acid comprising the murine  
 CC NR3 open reading frame cloned into an expression vector is  
 CC designated PLXSN.NNRG3 (ATCC 209156). The invention provides human  
 CC and murine NR3 polynucleotides (see also AAX06988) and polypeptides  
 CC (see also AAW97618), expression vectors, host cells and methods for  
 CC the recombinant production of novel NR3s. The polynucleotides and  
 CC polypeptides can be used to enhance the survival, proliferation or  
 CC differentiation of cells having the ErbB4 receptor in vivo and in  
 CC vitro. They can be used to prevent or treat damage to a nerve or  
 CC damage to other NR3-expressing or NR3-responsive cells, e.g.  
 CC brain, heart, or kidney cells. In particular, they can be used to  
 CC treat diseases which involve neural cell growth such as demyelination,  
 CC or damage or loss of glial cells (e.g. multiple sclerosis). They can  
 CC be used to treat patients whose nervous system has been damaged by  
 CC e.g. trauma, surgery, stroke, ischaemia, infection, metabolic  
 CC disease, nutritional deficiency, malignancy, or toxic agents. NR3  
 CC can also be used to treat motor neuron disorders such as amyotrophic  
 CC lateral sclerosis (Lou Gehrig's disease), Bell's palsy, and  
 CC conditions involving spinal muscular atrophy or paralysis. NR3

CC can be used to treat human neurodegenerative disorders such as  
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple  
CC sclerosis, Huntington's chorea, Down's syndrome, nerve deafness,  
CC and Meniere's disease. They can also be used to treat neuropathies  
CC associated with systemic disease including post-polio syndrome,  
CC hereditary neuropathies including Charcot-Marie-Tooth disease,  
CC Refsum's disease, abetalipoproteinemia, Ranganer disease, Krabbe's  
CC disease, metachromatic leukodystrophy, Fabry's disease and  
CC Dystonia-tortia syndrome, to treat disease of skeletal muscle of  
CC smooth muscle, such as muscular dystrophy or diseases caused by  
CC skeletal or smooth muscle wasting. The products can also be used  
CC for detection, diagnosis, for the production of transgenic or  
CC knockout animals or for drug screening.  
XX  
SQ Sequence 2538 BP; 606 A; 785 C; 649 G; 498 T; 0 other;

alignment\_scores:  
Quality: 277.00 Length: 47  
Ratio: 5.894 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-877-665-4 x AAX06987 ..

Align seg 1/1 to: AAX06987 from: 1 to: 2538

1 HisphelyProCyArgAspLysAspLeuAlaTyrCysLeuAsnAspG 17  
|||  
1150 CACTTCACACCTGCTGAGACACGACCTGCGTATTCCTCAAGATG 1199  
17 ylglyCyspheVal11egluThrLeuThrGlySerHisLysHisCysArg 34  
|||  
1200 TAAATGCTTGTGATTCAGACCTGACACCTGATCCATAGCACCTGCGT 1249  
34 ylglyGluGlyTyrGlnGlyValArgCysAspGlnPheLeu 47  
|||  
1250 GCAAGGAGGCTACCAAGAGATCGCTGTGATCAATTCTG 1290

seq\_name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX36424

seq\_documentation\_block:  
ID AAX36424 standard; DNA; 536 BP.  
XX  
AC AAX36424:  
XX  
DT 06-JUL-1999 (first entry)  
XX  
DE Human heregulin-like factor coding sequence clone HAGFE38R.  
XX  
KW Human heregulin-like factor; HLF; cell growth regulator; diagnosis;  
KW neural system disorder; cancer; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO9857989-A1.  
XX  
PD 23-DEC-1998.  
XX  
PF 16-JUN-1998; 98WO-US12403.  
XX  
PR 17-JUN-1997; 97US-0049942.  
XX  
PA (HDM-) HUMAN GENOME SCI INC.  
XX  
PI (GEOU) UNIV GEORGETOWN.  
XX  
PI H1Jaz1 MM, King CR, Ruben SM, Young P;  
XX  
XX WPI: 1999-095327/08.  
XX  
XX New isolated heregulin-like factor - used to develop products for  
XX the diagnosis and treatment of disorders involving regulation of  
XX cell growth, particularly cancers

XX  
XX Claim 20; Page 90-91; 118pp; English.  
XX  
CC This sequence encodes the human heregulin-like factor (HLF) of the  
CC invention. The HLF is involved in the regulation of cell growth.  
CC Detection of different levels of expression of the HLF gene can be used  
CC for the diagnosis of disorders, e.g. in the neural system. In  
CC particular, detection of different levels of HLF gene expression in cells  
CC or body fluid of an individual can be used for diagnosing cancer. The  
CC products can also be used in the treatment of disorders involving  
CC abnormal levels of HLF activity.  
XX  
SQ Sequence 536 BP; 141 A; 131 C; 125 G; 120 T; 19 other;

alignment\_scores:  
Quality: 251.00 Length: 49  
Ratio: 5.340 Gaps: 2  
Percent Similarity: 95.918 Percent Identity: 95.918

alignment\_block:  
US-09-877-665-4 x AAX36424 ..

Align seg 1/1 to: AAX36424 from: 1 to: 536

1 HisphelyProCyArgAspLysAspLeuAlaTyrCysLeuAsnAspG 17  
|||  
99 CACTTCACACCTGCGGAGACACCTGCGATGCTCATGATG 148  
17 ylglyCyspheVal11egluThrLeuThrGlySerHisLysHisCysArg 33  
|||  
149 GCGAGTCTTGTGATTCGAAACCCGACCGGATCCCATTAACACTGCG 198  
33 GCTGlyGluGlyTyrGlnGlyValArgCysAspGlnPheLeu 47  
|||  
199 GTGCAAGAGGCTACCAAGAGATCGCTGTGATCAATTCTG 241

seq\_name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA60998

seq\_documentation\_block:  
ID AAA60998 standard; CDNA; 156 BP.  
XX  
AC AAA60998:  
XX  
DT 09-NOV-2000 (first entry)  
XX  
DE Human NDF EGF-like domain derived peptide encoding CDNA SEQ ID NO.2.  
XX  
KW Human; sensory epithelial cell; growth; stimulant; inner ear; EGF;  
KW epithelial growth factor; NDF; heregulin; monoclonal antibody;  
KW adult rat utricular epithelium; ss.  
XX  
OS Homo sapiens.  
XX  
PN US6080845-A.  
XX  
PD 27-JUN-2000.  
XX  
PF 28-JAN-1999; 99US-0238182.  
XX  
PR 05-AUG-1998; 98US-0129549.  
XX  
PA (AMGE-) AMGEN INC.  
XX  
PI Carnehan JF;  
XX  
XX WPI: 2000-451229/39.  
XX  
XX P-PSDB; AAB12602.  
XX  
XX Novel monoclonal antibody against adult rat utricular epithelium useful  
XX in study and research with such organs and tissue  
XX  
XX Disclosure; Column 11-12; 12pp; English.



cc The glial ce

PS Disclosure; Flg 19; 102pp; English.

neuropathy etc.

Disclosure; Fig 19, 11pp; English.

AAT30989-T30994 represent the coding sequences for glial growth factor (GGF) peptides which encompass the epidermal growth factor-like domain (EGFL). These sequences can be used in the method of the invention.

This method is to affect cellular communication in a vertebrate by administration of neuroligin or a neuroligin-producing cell, where the neuroligin interacts with a first cell type to form product A, which in turn affects a function of a second cell type. The effect on the second cell type results in the production of product B which affects the first (or a third) cell type. Alternatively a neuroligin coding sequence is incorporated into the genome of a vertebrate cell, and the neuroligin is then expressed by the cell. The first cell type is a nervous system support cell, preferably a Schwann cell, neuron or muscle cell. The second and third cell types are nervous system or muscle cells. The products A and B are neurotrophic agents, matrix molecules, proteases or protease inhibitors. The method can be used to treat neurological disorders, such as peripheral neuropathy, amyotrophic lateral sclerosis, spinal muscular atrophy, nerve injury, Alzheimer's disease or spinal cord injury. It can also induce muscle mitogenesis. This method preferably induces more than one neurotrophic product, and is more effective than using a single neurotrophic factor, which would tend to have a restricted effect on a specific neuronal subtype.

Sequence 192 BP; 61 A; 39 C; 50 G; 42 T; 0 other;

Alignment scores:

	Quality: 113.50	Length: 46
Ratio:	3.338	Gaps: 1
Percent Similarity:	73.913	Percent Identity: 34.783

alignment\_block:  
US-09-877-665-4 x AAT30990 ..

Align seg 1/1 to: AAT30990 from: 1 to: 192

```
1 HspPhelyProGysArgGAspLysAspLeuAlaTyrCysLeuAAsnAspCl 17  
||||| || :|||::|::|::|::|::|::| | |  
4 CATCTGTGCAGAGTGTCGACAGAAAGGAATAACTTTCTGTGTGTAATGGAGG 53  
  
17 yGlucySPheValIleqlurThrLeuThrGlySerHisLysHis...CysA 33  
|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
54 CGAGCGCTCATGTGTAAGAAGCCATTCAATCCCTCAAGATACTTTGTGCA 103  
  
33 rGcysLysGLuGIlyTYrGlnGlyVALrGcysArSeln 45  
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
104 AGTGCACAACCTGGATTCACGTGAGCCGAGATGATCTGTAG 141
```

seq\_name: /STDS1/gcgdata/geneseq/geneseqn\_emb1/NA1996.DAT.AAT06734

seq\_documentation\_block:

ID	AAT06734 standard; DNA; 192 BP.
XX	AC
XX	AAT06734;
XX	XX
DE	Epidermal growth factor-like domain GGF peptide EGFL2 coding sequence.
XX	XX
KW	glial growth factor; GGF; human; hGGF2; Schwann cell; mitogenesis; GGF-1; neural cell; therapy; peripheral nerve damage; demyelination; bovine; neurodegenerative disorder; neural regeneration; acetylcholine receptor; fibroblast proliferation; wound repair; multiple sclerosis; glial tumour; GGF-II; epidermal growth factor; hEGF; ss.
OS	Synthetic.
XX	XX
PN	MO9532724-Al.
07-DEC-1995.	XX
DD	XX



KW epidermal growth factor; EGF; SS.

PD 17-FEB-1994.

XX	PN	MO9426298-A.
XX	PD	24-NOV-1994.
XX	XX	
XX	PF	06-MAY-1994; 94WO-US05083.
XX	PR	06-MAY-1993; 93US-0059022.
XX	PR	08-MAR-1994; 94US-0209204.
XX	PA	(CAME-) CAMBRIDGE NEUROSCIENCE.
XX	PI	Gwynne DI, Marchionni M, Sklar R;
XX	DR	WPI; 1995-006353/01.
XX	DR	P-PsDB; AAR67254.
XX	PT	Treating mammalian muscle diseases and disorders - by admin. of
XX	PT	GGF2 and other specified polypeptide(s) which bind the p185erbB2
XX	PT	receptor.
XX	PS	Claim 38; Page 151; 241pp; English.
XX	AAQ7919	encodes AAR67254 human epidermal like growth factor 6 (EGFL6);
XX	CC	The glial cell mitogenic activity of Egfr6 can be used to treat a
XX	CC	variety of mammalian skeletal, cardiac and smooth muscle diseases,
XX	CC	including acetylcholine receptor deficiency.
XX	Sequence	252 BP; 78 A; 54 C; 63 G; 57 T; 0 other;

```
alignment_scores:
  Quality: 113.50
  Ratio: 3.338
  Percent Similarity: 73.913
  Length: 46
  Gaps: 1
  Percent Identity: 34.783
```

alignment\_block:  
US-09-877-665-4 x AAO74919 .

Align seg 1/1 to: AAQ74919 from: 1 to: 252

```

1 HisPhelyProCysAsnArgpLysAspLeuAlaIleCysLeuAsnAspGln 17
   |||||  |||  :|||:|||||:|||||:|||||  ||
4 CACCTTGTCACAGTGCACAGAGAGAGAAACCTTCTGTGTGAAATGGAGG  53
   |||||:|||||:|||||:|||||:|||||:|||||
17 yGluCysPheAlaIleGluThrLeuThrGlySerHisLysHis...CysA  33
   |||||:|||||:|||||:|||||:|||||:|||||
54 CGAGTCTTCACAGGTGAGGAAAGACCTTCAAAATCCTCAAGAAATCTTGTGCA  103
   |||||:|||||:|||||:|||||:|||||:|||||
33 rGcysLysGluLysIleYrGlnGlyValaIArgCysAspGln 45
   ::|||::|||:|||||:|||||:|||||:|||||
104 AGTGCACCACTGTGATTACTGTGAGGAGGAATGATTAATCTAGAG  141

```

**THIS PAGE BLANK (USPTO)**

---

OM of: US-09-877-665-4 to: EST:\* out\_format : pfs  
Date: Sep 15, 2002 12:55 PM

About: Results were produced by the Gencore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

## Command line parameters:

-MODEL=frame-p2n.model -DEV=x1h  
-Q/cgn2.1/USPTO-spool/US09877665/runat.13092002.083912.15645/app-query.fasta.1.103  
-DB=EST -FASTA=fastap -SUFFIX=p2n.rst -GAPOP=12.000 -CAPEXT=4.000  
-MINMATCH=0.100 -LOCPCL=0.000 -LOOPEXT=0.000 -OGAPOP=4.500  
-OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000  
-RGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000  
-DELEXT=7.000 -START=1 -MATRIX=blissum62 -TRANS=human40.cdi  
-LIST=45 -DCCALIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0  
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500  
-MINLEN=0 -MAXLEN=2000000000 -USER=US09877665.GCEN1.1.2394  
-NCPU=6 -ICPU=3 -LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30  
-NO\_XLPHY -WAIT -THREADS=1

## Search information block:

Query: US-09-877-665-4

Query length: 47

Database: EST\*

Database sequences: 13736207

Database length: 1841575050

Search time (sec): 1760.310000

## Score list:

Sequence	Strd Orig	ZScore	EScore	Len	Document
gb_est1:AL534571	271.00	615.80	4.7e-25	785	AL534571 AL534571 LTI_FL013_FBR
gb_est2:CNS05BVV	167.00	371.93	1.8e-11	1001	AL330340 Tetradon nigroviridis
gb_est2:BM142279	113.50	253.45	7.2e-05	570	BM142279 lf33c10.y1 Melton North
gb_est2:BF050411	113.50	249.21	0.0001	866	BP030411 601558480F1 NIH_MGC_58
gb_est2:BG150472	111.50	247.60	0.0002	642	BG150472 7K01e09.x1 NCI_CGAP_GC
gb_est2:BP937937	110.50	247.61	0.0002	510	BP937937 fm70e01.y1 NCI_CGAP_GC
gb_est1:AA238077	110.00	245.36	0.0002	568	AA238077 mx10e03.y1 Soares mous
gb_est1:BB637211	110.00	243.96	0.0002	652	BB637211 BB637211 RIKEN full-16
gb_est2:BI908144	104.50	231.12	0.0013	657	BI908144 603067448F1 NIH_MGC_11
gb_est2:BI907799	104.50	228.14	0.0018	881	BI907799 603065533F1 NIH_MGC_11
gb_hic:BC006492	104.50	221.53	0.0043	1692	BC006492 Homo sapiens: Angelin
gb_est1:AA968077	104.00	237.03	0.0006	327	AA968077 un03h01.r1 Soares mous
gb_est2:BE648742	103.50	235.01	0.0008	336	BE648742 UT-M-BR2.2-acl-d-09-0-
gb_est1:AI1197081	103.50	233.50	0.0009	413	AI1197081 ud5fco2.r1 Soares-NMPu
gb_est1:AA798243	103.50	231.50	0.0012	503	AA798243 vy05e02.r1 Stratiogene
gb_est2:BF941240	101.00	232.43	0.0013	526	BF941240 7d95f08.x1 Lupski dors
gb_est2:BE648780	101.00	232.43	0.0011	259	BE648780 UT-M-BR2.2-aop-b-12-0-
gb_est1:AV851999	97.00	216.97	0.0077	476	AV851999 AV851999 Nori Satoh un
gb_est2:BI464093	97.00	210.67	0.0173	887	BI464093 603302870F1 NIH_MGC_97
gb_est2:BG213584	96.50	211.99	0.0146	694	BG213584 RST31316 Athysys RAGH
gb_est2:BG195532	95.50	211.90	0.0148	557	BG195532 RST31316 Athysys RAGH
gb_est1:BB381556	95.50	210.78	0.0171	622	BB381556 BB381556 RIKEN full-16
gb_gss:CNS0272C	95.50	207.36	0.0265	872	AL213825 Tetradon nigroviridis
gb_est1:BB489119	93.50	213.93	0.0119	296	BB489119 BB489119 RIKEN full-16
gb_est1:AV904635	93.00	203.93	0.0412	690	AV904635 AV904635 Nori Satoh un
gb_est2:BG196928	93.00	201.60	0.0555	868	BG196928 RST16159 Athysys RAGH
gb_est2:BG197944	91.50	212.80	0.0132	204	BG197944 RST16159 Athysys RAGH
gb_est1:AV878992	91.00	201.22	0.0582	570	AV878992 AV878992 Nori Satoh un
gb_est1:BB489119	91.00	200.13	0.0670	635	BB489119 BB489119 RIKEN full-16
gb_est1:AM762061	90.00	207.01	0.0277	256	AM762061 uc33c01.y1 NCI_CGAP_MG
gb_est1:AM76657	90.00	199.66	0.0712	529	AM76657 uc79e01.y1 NCI_CGAP_MG
gb_est2:BI6151936	90.00	195.53	0.1208	795	BI6151936 603298677F1 NIH_CGAP_L
gb_est2:BI410828	90.00	193.73	0.1521	949	BI410828 602995334F1 NCI_CGAP_L
gb_est1:AV897556	89.00	193.60	0.1548	765	AV897556 AV897556 Nori Satoh un
gb_est1:AV853430	89.00	193.23	0.1622	793	AV853430 AV853430 Nori Satoh un
gb_est2:BG343733	88.00	198.28	0.3062	1028	BG343733 HSM5E0006016F Hordeu
gb_est1:AI478167	87.50	192.08	0.1880	630	AI478167 tm44h09.x1 NCI_CGAP_K
gb_est2:BI145782	87.00	191.83	0.1942	576	BI145782 602909246F1 NCI_CGAP_L
gb_est1:AV616555	87.00	191.30	0.2079	607	AV616555 AV616555 Bos taurus ov
gb_est1:AA990660	87.00	190.48	0.2309	658	AA990660 LD344470.5pTime LD Dros

gb_est1:BB646928	+	87.00	189.77	0.2530	706	BB646928 BB646928 RIKEN full1
gb_est2:BG261805	+	87.00	189.61	0.2581	717	BG261805 602373601F1 NIH_MGC
gb_est2:BI330162	+	87.00	189.17	0.2732	749	BI330162 602983780F1 NCI_CGA
gb_est2:BF237257	+	87.00	188.77	0.2875	779	BF237257 602025418F1 NCI_CGA
gb_est2:BI148137	+	87.00	188.20	0.3093	824	BI148137 602912309F1 NCI_CGA

seq\_name: gb\_est1:AL534571

## seq\_documentation\_block:

LOCUS AL534571 785 bp mRNA linear EST 13-FEB-2001  
DEFINITION AL534571 LTI\_FL013\_FBRn1 Homo sapiens CDNA clone CS0DF0041618 5  
PRIME, mRNA sequence.  
ACCESSION AL534571  
VERSION AL534571.1 GI:12798064  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 785)

Li, W.B., Gruber, C., Jessee, J., and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

## COMMENT

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr.

## FEATURES

source

1..785

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CS0DF0041618"

/dev\_stage="pooled tissue from post conception fetuses (20

week, 24 week and 26 week)"

/lab\_host="DH10B"

/note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand

cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-stranded cDNA was digested with Not I

and cloned into the Not I and Eco RV sites of the

pCMVSPORT 6 vector. Library was constructed by Life

Technologies. Contact: Feng Liang Life Technologies, a

division of Invitrogen 9800 Medical Center Drive Rockville

, Maryland 20850, USA Fax: (1) 301 610 8371 Email:

liang@lifetech.com URL:

http://fulllength.invitrogen.com" 3 others

## BASE COUNT

222 a 161 c 181 g 218 t

## ORIGIN

1

HisphelysProcysa

atgagp

aspleu

atg

atg

atg

atg

atg

atg

atg

atg

atg

atg

atg

atg

atg

atg

atg

atg

atg

atg

atg

atg

atg

atg

atg

atg

atg

atg

atg

```

seq_documentation_block:
LOCUS      CNS05BVY                      1001 bp    DNA      linear      GSS 26-MAY-2000
DEFINITION Tetracodon nigroviridis genome survey sequence T3 end of clone
            012D17 of library C from Tetracodon nigroviridis, genomic survey
            sequence.
ACCESSION  AL330340.1 GI:8223962
VERSION    AL330340
KEYWORDS   GSS: genome survey sequence.
SOURCE     Tetracodon nigroviridis.
            Tetracodon nigroviridis.
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetracodontiformes;
            Tetracodontidae; Tetracodon.
REFERENCE  1 (bases 1 to 1001)
            Roest-Crollius,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
            Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetracodon nigroviridis
            Unpublished
            2 (bases 1 to 1001)
            Roest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
            Human gene number estimate provided by genome wide analysis using
            Tetracodon nigroviridis DNA sequence
            Unpublished
            3 (bases 1 to 1001)
            Genoscope.
            Direct Submission
            Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
            This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetracodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetracodon.
FEATURES
    source
        1..1001
            /organism="Tetracodon nigroviridis"
            /db_xref="taxon:99883"
            /clone_1lb="012D17"
            /clone_1lb="C"
            /note="Genoscope sequence ID : C0AC012CB09T1-end : T3"
BASE COUNT      272 a      218 c      219 g      278 t      14 others
ORIGIN
alignment_scores:
    Quality: 167.00      Length: 33
    Ratio: 5.387      Gaps: 0
    Percent Similarity: 93.939      Percent Identity: 81.818
alignment_block:
US-09-877-665-4 x CNS05BVY
Align seg 1/1 to: CNS05BVY from: 1 to: 1001

1 HisphelysProcysargaspLysaspLeuAlaTyrCysLeuAsnAspG1 17
|||||
667 CATTTCAAGCCCTGCATGAGAGACCTGCGTACTGTGTGAGACGATGG 716
|||||
17 yglucySpheValIleGIuThrLeuThrGlySerHisLysHisGysArg 33
|||||
717 AGAGTGTTCCTCATCGAGAGCTCAGCGGCGCTCACAAACACTGCAG 765
|||||
seq_name: gb_est2:BM142279
seq_documentation_block:
LOCUS      BM142279                      570 bp    mRNA      linear      EST 29-NOV-2001
DEFINITION t133c10.y1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
            CDNA 5' similar to TR:035947 035947 PRO-NEUREGULIN-1 PRECURSOR ;,
            mRNA sequence.

```

```

ACCESSION  BM142279                      GI:17152346
VERSION    BM142279.1
KEYWORDS   EST.
SOURCE     human.
            Homo sapiens
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 570)
            Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Maira,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritzer,E., Ronko,I., Bennett,D., Cardenas
            M., Gibbons,M., McCann,R., Cole,R., Tsagaratshvilli,R., Williams,T.,
            Jackson,X. and Bowers,Y.
            Endocrine Pancreas Consortium
            Unpublished (2000)
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Douglas Melton DNA sequencing by:
            Washington University Genome Sequencing Center for information on
            obtaining a clone please contact: Juliana Brown
            (brownjfas.harvard.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 429.
FEATURES
    source
        1..570
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_1lb="Melton Normalized Human Islet 4 N4-HIS 1"
            /sex="Both"
            /tissue.type="Islets of Langerhans"
            /dev_stage="Adult"
            /lab_host="DH10B"
            /note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
            Site_2: Sal 1; Starting library constructed using
            Superscript plasmid library kit (Life Technologies). cDNA
            made by oligo-dt priming. Size-selected by column
            fractionation; average insert size 1.08 kb. Library was
            amplified once on solid support and plasmid DNA from
            library was prepared. The library DNA was normalized by
            method #4 from Bonaldo, Lennon, and Soares 1996 Genome
            Research 6:791-806; 0.5 microgram single-stranded library
            plasmid DNA was mixed with 5 micrograms PCR product
            representing library inserts and hybridized to an EcoT of
            20. Single-stranded (unhybridized) plasmids were isolated
            by hydroxyapatite chromatography and used to make this
            library."
BASE COUNT      175 a      140 c      132 g      123 t
ORIGIN
alignment_scores:
    Quality: 113.50      Length: 46
    Ratio: 3.338      Gaps: 1
    Percent Similarity: 73.913      Percent Identity: 34.783
alignment_block:
US-09-877-665-4 x BM142279
Align seg 1/1 to: BM142279 from: 1 to: 570

1 HisphelysProcysargaspLysaspLeuAlaTyrCysLeuAsnAspG1 17
|||||
157 CATCTTTAAATGTGCGGAGAGAGAAAACCTTCTGTGTGAGATGAGG 206
|||||
17 yglucySpheValIleGIuThrLeuThrGlySerHisLysHis...CysA 33
|||||

```



```

17 yglucysphevalilleglthleuthrlyserhislvs...Cysa 33
|||||.....:|||||
505 GGAGTCTTCATGCTGAAGACCTTCAACCCCTGAGATCTGTGCA 456
33 rgcylslysgluclytyrtinglyvalarvcysaspcln 45
:::|||||.....:|||||
455 AGTGCACACCTGATCTGAGGACAGATGACTGAG 418

seq_name: gb_est1:BF937937

seq_documentation_block:
LOCUS BF937937 510 bp mRNA linear EST 13-FEB-2002
DEFINITION f070601.y1 zebrafish adult retina cDNA Danio rerio cDNA clone
4200673.5 similar to TR:005199 Q05199 PRO-NEUREGULIN-1 ; mRNA
sequence.
ACCESSION BF937937
VERSION BF937937.1 GI:12355257
KEYWORDS EST.
SOURCE zebrafish.
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
; Cyprinidae; Danio.
1 (bases 1 to 510)
REFERENCE
AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
,R., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
and Wilson,R.
Mashu zebrafish EST Project 1998
JOURNAL Unpublished (1998)
COMMENT Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@wustl.edu
Library constructed by: Chandra Tucker and Gregory Niemi DNA
Sequencing by: Washington University Genome Sequencing Center Clone
distribution: Ressourcenzentrum für Primatdatenbank, Berlin, Germany
(web address: www.rzpd.de)
Seq primer: T3 ET from Amersham
High quality sequence stop: 419.
Location/Qualifiers
FEATURES
Source 1..510
/organism="Danio rerio"
/strain="wild-type"
/db_xref="taxon:7955"
/clone="4200673"
/clone_1lb="zebrafish adult retina cDNA"
/sex="mixed"
/dev_stage="1-2 years"
/lab_host="E.Coli XL1-Blue MRF' (XL1-Blue MRF')
"
/notes="Vector: Lambda ZAP II (pbluescript SK-); Site:1;
EcoRI; Site:2; SalI; This zebrafish library was
constructed by Dr. Susan E. Brockerhoff (email:
sbrocker@eu.washington.edu) RZPD library number: 760"
BASE COUNT 143 a 122 c 125 g 119 t 1 others
ORIGIN
alignment_scores:
Quality: 110.50 Length: 48
Ratio: 3.157 Gaps: 1
Percent Similarity: 72.917 Percent Identity: 35.417

alignment_block:
US-09-877-665-4 x BF937937 ..
Align seg 1/1 to: BF937937 from: 1 to: 510
1 HisphelysProcysatrgasplysaspLeuAlatyrcysLeuasnasp1 17

```

```

328 CACGGAGACCCCTGCACAGACGACGAGAAAGATAGTGTGTAACCATG 377
|||||.....:|||||
17 yglucysphevalilleglthleuthrlyser...HislvsHscysa 33
|||||.....:|||||
378 CAAGTCTTCACGCTGAGGTGCACACCCGCAACATCAGACGCCCTTGCA 427
|||||.....:|||||
428 GGTGCCCAATGATGTTACTGCTGATCGCTGCCAACACTAGCTA 471

seq_name: gb_est1:AA238077

seq_documentation_block:
LOCUS AA238077 568 bp mRNA linear EST 03-MAR-1997
DEFINITION Mx10603.r1 Soares mouse NML Mus musculus cDNA clone IMAGE:679804
5', mRNA sequence.
ACCESSION AA238077
VERSION AA238077.1 GI:1862116
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 568)
REFERENCE
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,R., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The Mashu-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
Mashu-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@lml.lnl.gov) for further information.
MGI:419508
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 459.
Location/Qualifiers
FEATURES
Source 1..568
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:679804"
/clone_1lb="Soares mouse NML"
/tissue_type="liver"
/lab_host="DH10B"
/notes="Vector: pT773D-Pac (Pharmacia) with a modified
polylinker; Site:1; Not I; Site:2; Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTCACATCGAGAGTGGAGCGCGCGAGATCTTTTCTTTTCTTTT 3'];
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT773 vector. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."
BASE COUNT 147 a 151 c 129 g 141 t
ORIGIN
alignment_scores:
Quality: 110.00 Length: 45
Ratio: 3.667 Gaps: 1
Percent Similarity: 66.667 Percent Identity: 42.222

alignment_block:
US-09-877-665-4 x AA238077 ..

```

FEATURES	Location/Qualifiers
source	1. .652

**Tissue Procurement:** Life Technologies, Inc.  
**CDNA Library Preparation:** Life Technologies, Inc





WashU-HHMI Mouse EST Project

discovery  
Genome Res. 6 (9): 791-806 (1996)

**MEDLINE** 97044477

## COMMENT

Contact: Chin, H  
National Institute of Mental Health  
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
20892-9643, USA  
Tel: 301 443 1706  
Fax: 301 443 9890  
Email: mestr@mail.nih.gov

CDNA Library Preparation: M.B. Soares Lab Clone distribution:  
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It  
should be noted that Bento Soares is generating a small number of  
additional specialized non-redundant arrays of BMAP cDNAs whose  
availability will be considered under appropriate and limited  
collaborative arrangements

Seq primer: M13 Reverse.

## FEATURES

Source Location/Qualifiers

1..356  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UT-M-BH2.2-aol-d-09-0-UT"  
/dev\_stage="27-32 days"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; The  
NIH-BMAP\_M.S3.2 library is a subtracted library of a  
series, ultimately derived from a mixture of individually  
tagged normalized libraries from ten regions of the mouse  
brain (cerebellum, brain stems, olfactory bulbs,  
hypothalamus, cortex, amygdala, basal ganglia, pineal  
gland, striatum, hippocampus) after a series of  
subtractions to reduce the representation of cDNAs from  
which ESTs had already been generated. The following  
serially subtracted libraries were generated in this  
process: NIH-BMAP\_M.S3.2, NIH-BMAP\_M.S2, NIH-BMAP\_M.S1.  
The subtracted library (NIH-BMAP\_M.S3.2) was constructed  
as follows: PCR amplified cDNA inserts from NIH-BMAP\_M.S2  
clones from which 3' ESTs had been derived was used as a  
driver in a hybridization with the NIH-BMAP\_M.S2 library  
in the form of single-stranded circles. The remaining  
single-stranded circles (subtracted library) was purified  
by hydroxyapatite column chromatography, converted to  
double-stranded circles and electroporated into DH10B  
bacteria (Life Technologies) to generate the  
NIH-BMAP\_M.S3.2 library. This procedure has been  
previously described (Bonaldo, Lennon and Soares, Genome  
Research 6:791-806, 1996)."

BASE COUNT

97 a 106 c 74 g 79 t

## alignment\_scores:

Quality: 103.50 Length: 48  
Ratio: 2.957 Gaps: 1  
Percent Similarity: 72.917 Percent Identity: 31.250

## alignment\_block:

US-09-877-665-4 x BE648742 ..

Align seg 1/1 to: BE648742 from: 1 to: 356

```

1 HspHelyProCySarGAspLysAspLeuAlaTyrCysLeuAsnAspG1 17
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
187 CATCTCATTAAGAGTGTGGGAGAGGAAACCTTCTGTGTGAATGAGG 236
17 yGluCysPheValIleGluThrLeuThrglySerHisLysHis..CysA 33
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
237 CGAGTCTTCATGCTGAAGACCTGTCAAAACCCCTCAAGTACTTGTGCA 286
33 rGcysLysGluGlyTyrGlnGlyValAlaGlyCysAspGlnPheLeu 47
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
287 AGTGCCTTCATGCTGAAGACCTGTCAAAACCCCTCAAGTACTTGTGCA 330

```

seq\_name: gb\_est1:A1197081

seq\_documentation\_block:

LOCUS A1197081

413 bp mRNA linear EST 14-OCT-1998

DEFINITION uc57c02.t1 Soares\_NMPu Mus musculus cDNA clone IMAGE:144986 5' similar to TR:007111 007111 GLIAL GROWTH FACTOR 2 PRECURSOR ;, mRNA

ACCESSION A1197081

GI:3749687

VERSION A1197081

EST

KEYWORDS

house mouse.

SOURCE

house mouse.

ORGANISM

house mouse.

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.

Warr, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,

Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,

Thelning, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and

Waterston, R.

The WashU-HMI Mouse EST Project

TITLE

Unpublished (1996)

JOURNAL

Contact: Marra M/Mouse EST Project

COMMENT

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:923302

Seq primer: -28m13 rev2 ET from Amerisham

High quality sequence stop: 412.

FEATURES

Location/Qualifiers

source

1..413

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="IMAGE:144986"

/clone\_lib="Soares\_NMPu"

/sex="female"

/dev\_stage="adult"

/lab\_host="DH10B"

/note="Organ: uterus; Vector: pT73D-Pac (Pharmacia) with

a modified polylinker; 1st strand cDNA was prepared from

pregnant mouse uterus, and was then primed with a Not I-

oligo(dT) primer. Double-stranded cDNA was ligated to Eco

RI adaptors (Pharmacia), digested with Not I and cloned

into the Not I and Eco RI sites of the modified pT73

vector. Library is normalized. Library was constructed by

Bento Soares and M. Fatima Bonaldo."

BASE COUNT

110 a 103 c 91 g 109 t

ORIGIN

```

1 HspHelyProCySarGAspLysAspLeuAlaTyrCysLeuAsnAspG1 17
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
168 CATCTCATTAAGAGTGTGGGAGAGGAAACCTTCTGTGTGAATGAGG 217
17 yGluCysPheValIleGluThrLeuThrglySerHisLysHis..CysA 33
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
218 CGAGTCTTCATGCTGAAGACCTGTCAAAACCCCTCAAGTACTTGTGCA 267

```

## alignment\_scores:

Quality: 103.50 Length: 48  
Ratio: 2.957 Gaps: 1  
Percent Similarity: 72.917 Percent Identity: 31.250

## alignment\_block:

US-09-877-665-4 x A1197081 ..

Align seg 1/1 to: A1197081 from: 1 to: 413

```

1 HspHelyProCySarGAspLysAspLeuAlaTyrCysLeuAsnAspG1 17
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
168 CATCTCATTAAGAGTGTGGGAGAGGAAACCTTCTGTGTGAATGAGG 217
17 yGluCysPheValIleGluThrLeuThrglySerHisLysHis..CysA 33
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
218 CGAGTCTTCATGCTGAAGACCTGTCAAAACCCCTCAAGTACTTGTGCA 267

```



**THIS PAGE BLANK (USPTO)**

---

OM of: US-09-877-665-4 to: GenEmbl: \* out\_format : pfs

Date: Sep 15, 2002 1:28 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL-frame+;p2n,model -DEV-x1h  
-O/-cgn2.1/USPto.spool/US09877665/Runat\_13092002\_083912\_15659/app-query.fasta.1.103  
-DB-GenEmbl -OPMT-fastap -SUFFIX-p2n.rge -GAPOP-12.000  
-GAPEXT-4.000 -MINMATCH-0.100 -LOOPEXT-0.000  
-GAPOP-4.500 -GAPEXT-0.050 -XGAPOP-10.000 -XGAPEXT-0.500  
-XGAPOP-6.000 -XGAPEXT-7.000 -XGAPOP-10.000 -XGAPEXT-0.500  
-DELOP-6.000 -DELEXT-7.000 -START-1 -MATRIX-biosum62  
-TRANS-numa40.cdi -LIST-45 -DOCALLIGN-200 -THR\_SCORE-pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN-15 -MODE-LOCAL -OUTFMT-pfs  
-NOR-ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09877665.cgn1.1.3708 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPRY -WAIT -THREADS=1

Search information block:

Query: US-09-877-665-4

Query length: 47

Database: GenEmbl: \*

Database length: 187333701

Search time (sec): 1997.040000

Score list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
gb_ro:AF010130	277.00	669.10	6.6e-29	2538	AF010130 Mus musculus neuregulin-3 (NRG3) mRNA, complete cds.
gb_hlg:AC096527	203.00	450.32	1.0e-16	137600	AC096527 Rattus norvegicus c
gb_pr:AL391478	195.00	427.98	1.8e-15	180625	AL391478 Human DNA sequence
gb_ov:AF076618	126.50	289.24	9.5e-08	3105	AF076618 Xenopus laevis neureg
gb_pat:AR025501	116.50	288.69	1.0e-07	156	AR025501 Sequence 2 from paten
gb_pat:AR068979	113.50	279.44	3.3e-07	192	AR068979 Sequence 155 from pat
gb_pat:AR010335	113.50	279.44	3.3e-07	192	AR010335 Sequence 155 from pat
gb_pat:AR010335	113.50	279.44	3.3e-07	192	AR010335 Sequence 155 from pat
gb_pat:AR133506	113.50	279.44	3.3e-07	192	AR133506 Sequence 155 from pat
gb_pat:AR141835	113.50	279.44	3.3e-07	192	AR141835 Sequence 155 from pat
gb_pat:AR143352	113.50	279.44	3.3e-07	192	AR143352 Sequence 155 from pat
gb_pat:AR151961	113.50	279.44	3.3e-07	192	AR151961 Sequence 155 from pat
gb_pat:136441	113.50	279.44	3.3e-07	192	136441 Sequence 155 from pat
gb_pat:188141	113.50	279.44	3.3e-07	192	188141 Sequence 155 from pat
gb_pat:AR022505	113.50	277.21	4.5e-07	252	AR022505 Sequence 159 from pat
gb_pat:AR068983	113.50	277.21	4.5e-07	252	AR068983 Sequence 159 from pat
gb_pat:AR103363	113.50	277.21	4.5e-07	252	AR103363 Sequence 69 from pat
gb_pat:AR133501	113.50	277.21	4.5e-07	252	AR133501 Sequence 159 from pat
gb_pat:AR141839	113.50	277.21	4.5e-07	252	AR141839 Sequence 159 from pat
gb_pat:AR143356	113.50	277.21	4.5e-07	252	AR143356 Sequence 159 from pat
gb_pat:AR151965	113.50	277.21	4.5e-07	252	AR151965 Sequence 159 from pat
gb_pat:136445	113.50	277.21	4.5e-07	252	136445 Sequence 159 from pat
gb_pat:188145	113.50	277.21	4.5e-07	252	188145 Sequence 159 from pat
gb_pat:AR022504	113.50	276.73	4.7e-07	267	AR022504 Sequence 158 from pat
gb_pat:AR068982	113.50	276.73	4.7e-07	267	AR068982 Sequence 158 from pat
gb_pat:AR103362	113.50	276.73	4.7e-07	267	AR103362 Sequence 67 from pat
gb_pat:AR133509	113.50	276.73	4.7e-07	267	AR133509 Sequence 158 from pat
gb_pat:AR141838	113.50	276.73	4.7e-07	267	AR141838 Sequence 158 from pat
gb_pat:AR143355	113.50	276.73	4.7e-07	267	AR143355 Sequence 158 from pat
gb_pat:AR151964	113.50	276.73	4.7e-07	267	AR151964 Sequence 158 from pat
gb_pat:136444	113.50	276.73	4.7e-07	267	136444 Sequence 158 from pat
gb_pat:188144	113.50	276.73	4.7e-07	267	188144 Sequence 158 from pat
gb_pr:HS02327	113.50	266.87	1.7e-06	886	HS02327 Human clone nbf19 neu d
gb_pat:AR022498	113.50	264.80	2.2e-06	1140	AR022498 Sequence 149 from pat
gb_pat:AR068976	113.50	264.80	2.2e-06	1140	AR068976 Sequence 149 from pat
gb_pat:AR103355	113.50	264.80	2.2e-06	1140	AR103355 Sequence 55 from pat
gb_pat:AR133503	113.50	264.80	2.2e-06	1140	AR133503 Sequence 149 from pat
gb_pat:AR141832	113.50	264.80	2.2e-06	1140	AR141832 Sequence 149 from pat
gb_pat:AR143349	113.50	264.80	2.2e-06	1140	AR143349 Sequence 149 from pat
gb_pat:AR151958	113.50	264.80	2.2e-06	1140	AR151958 Sequence 149 from pat

gb\_pat:136438 + 113.50 264.80 2.2e-06 1140 136438 Sequence 149 from pa  
gb\_pat:188138 + 113.50 264.80 2.2e-06 1140 188138 Sequence 149 from pa  
gb\_pat:AR29423 + 113.50 264.43 2.3e-06 1193 A29423 putative bovine GGF-  
gb\_pat:AR022483 + 113.50 264.43 2.3e-06 1193 AR022483 Sequence 134 from  
gb\_pat:AR068961 + 113.50 264.43 2.3e-06 1193 AR068961 Sequence 134 from

seq\_name: gb\_ro:AF010130

seq\_documentation\_block:

LOCUS AF010130 2538 bp mRNA linear ROD 23-SEP-1997  
DEFINITION Mus musculus neuregulin-3 (NRG3) mRNA, complete cds.  
ACCESSION AF010130  
VERSION AF010130.1 GI:2429163  
KEYWORDS  
SOURCE house mouse.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 2538)  
Zhang,D., Sliwkowski,M.X., Mark,M., Frantz,G., Akita,R., Sun,Y.,  
Hillman,K., Crowley,C., Brush,J. and Godowski,P.J.  
Neuregulin-3 (NRG3): A novel neural tissue-enriched protein that  
binds and activates ErbB4  
Proc. Natl. Acad. Sci. U.S.A. 94 (18), 9562-9567 (1997)

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

1 HisphenylprocyasargaspysasplealatyCysleuansapgl 17  
|||||

alignment\_scores:  
Quality: 277.00 Length: 47  
Ratio: 5.894 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

US-09-877-665-4 x AF010130  
Align seg 1/1 to: AF010130 from: 1 to: 2538

1150 CACTCAACCTGTCGACACAGACCTGGCGTATGTCTCATGATG 1199  
 17 yglucysphevalillegluThrleuThrlyserHislySHScysArgC 34  
 1200 TGAATGCTTGTGATGAGACCCCTGACAGATCCATTAAGCACTGCGGT 1249  
 34 yslsgsluGlutylTrgInglValAlaArgcysAspGlnPheIeu 47  
 1250 GCAAGGAAGGCTACCAAGAGTCCGTGTGATCAATTCTGTG 1290

seq\_name: gb\_htg:AC096527

seq\_documentation\_block:  
 LOCUS AC096527 137600 bp DNA linear HTG 20-DEC-2001  
 DEFINITION Rattus norvegicus clone CH230-78H3, \*\*\* SEQUENCING IN PROGRESS \*\*\*  
 67 unordered pieces.  
 AC096527  
 AC096527 GI:17962756  
 VERSION  
 KEYWORDS  
 SOURCE HTG; HTGS\_PHASE1.  
 ORGANISM  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus

REFERENCE  
 AUTHORS  
 1 (bases 1 to 137600)  
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
 Alsbrooks,S.L., Amaralunga,H.C., Are,J.R., Banks,T., Barbara,J.,  
 Benton,J., Bimaga,K., Blankenburg,K., Bohn,D., Bouck,J.,  
 Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Bunay,C.,  
 Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Cartron,T.F.,  
 Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,  
 Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C.,  
 Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,  
 Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,  
 Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Drepper,H.,  
 Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,  
 Elina,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,  
 Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,  
 Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,  
 Hamilton,K., Harris,C., Harris,K., Hart,M., Haylak,P., Hayes,A.,  
 Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,  
 Hollins,B., Homs,F., Howard,S., Huber,J., Huliy,S., Hume,J.,  
 Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jollivet,S.,  
 Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Kovach,J.,  
 Kovach,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,  
 Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,  
 Lounsged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,  
 Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,  
 Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M.,  
 Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,  
 Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,  
 Nguyen,A., Nguyen,N., Nguyen,M., Nickerson,E., Nwokenkwo,S.,  
 Ogih,M., Okunou,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,  
 Peery,U., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,  
 Quiles,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,  
 Slisdon,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H.,  
 Stone,H., Sutton,A., Sytek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,  
 Tang,H., Tansey,J., Taylor,C., Taylor,T., Tellrod,B., Thomas,N.,  
 Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,  
 Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,  
 Watlington,S., Williams,G., Williamson,A., Wleczkyk,R., Woodson,S.,  
 Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorilla,S., Nelson,D.,  
 Weinstein,G. and Gibbs,R.

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

2 (bases 1 to 137600)  
 Direct Submission  
 Submitted (18-SEP-2001) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Dec 20, 2001 this sequence version replaced gl.15638976.

----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information  
 Center project name: GCCT  
 Center clone name: CH230-78H3  
 ----- Summary Statistics  
 Assembly program: Phrap; version 0.990329first call to  
 findPhrapList  
 Consensus quality: 105498 bases at least Q40  
 Consensus quality: 121268 bases at least Q30  
 Consensus quality: 130770 bases at least Q20  
 Estimated insert size: 102221; sum-of-coverage  
 Quality coverage: 0x in Q20 bases; agarose-gel estimation  
 Quality coverage: 1.5x in Q20 bases; sum-of-coverage estimation

NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 NOTE: This is a 'working draft' sequence. It currently  
 consists of 67 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.

1	8775:	contig of 8775 bp in length
8776	8875:	gap of unknown length
8876	15580:	contig of 6705 bp in length
15581	15680:	gap of unknown length
15681	19503:	contig of 3823 bp in length
19504	19603:	gap of unknown length
19604	23557:	contig of 3954 bp in length
23558	23657:	gap of unknown length
23658	27350:	contig of 3693 bp in length
27351	27450:	gap of unknown length
27451	29673:	contig of 2223 bp in length
29674	29773:	gap of unknown length
29774	32280:	contig of 2507 bp in length
32281	32380:	gap of unknown length
32381	36076:	contig of 3696 bp in length
36077	36176:	gap of unknown length
36177	40285:	contig of 4109 bp in length
40286	40385:	gap of unknown length
40386	43380:	contig of 2995 bp in length
43381	43480:	gap of unknown length
43481	46160:	contig of 2680 bp in length
46161	46260:	gap of unknown length
46261	48577:	contig of 2317 bp in length
48578	48677:	gap of unknown length
48678	50498:	contig of 1821 bp in length
50499	50598:	gap of unknown length
50599	51940:	contig of 1342 bp in length
51941	52040:	gap of unknown length
52041	54297:	contig of 2257 bp in length
54298	54397:	gap of unknown length
54398	56690:	contig of 2293 bp in length
56691	56790:	gap of unknown length
56791	59221:	contig of 2431 bp in length
59222	59321:	gap of unknown length
59322	61180:	contig of 1859 bp in length
61181	61280:	gap of unknown length
61281	63486:	contig of 2206 bp in length
63487	63586:	gap of unknown length
63587	65586:	contig of 2000 bp in length
65587	65686:	gap of unknown length
65687	67367:	contig of 1681 bp in length
67368	67467:	gap of unknown length
67468	69491:	contig of 2024 bp in length
69492	71008:	gap of unknown length
71009		contig of 1417 bp in length
69592		

```

* 71009 71108: gap of unknown length
* 71109 72408: contig of 1300 bp in length
* 72409 72508: gap of unknown length
* 72509 73781: contig of 1273 bp in length
* 73782 73881: gap of unknown length
* 73882 75196: contig of 1315 bp in length
* 75197 75297: gap of unknown length
* 75297 77054: contig of 1758 bp in length
* 77055 77154: gap of unknown length
* 77155 78785: contig of 1631 bp in length
* 78786 78885: gap of unknown length
* 78886 80624: contig of 1738 bp in length
* 80624 82279: gap of unknown length
* 82279 82378: contig of 1555 bp in length
* 82379 84106: gap of unknown length
* 84107 84206: contig of 1728 bp in length
* 84207 85530: contig of 1324 bp in length
* 85531 87347: gap of unknown length
* 87348 87447: contig of 1717 bp in length
* 87448 88988: gap of unknown length
* 88989 90225: contig of 1541 bp in length
* 90226 90325: gap of unknown length
* 90326 91776: contig of 1137 bp in length
* 91777 91876: gap of unknown length
* 91877 93278: contig of 1402 bp in length
* 93279 93378: gap of unknown length
* 93379 94793: contig of 1415 bp in length
* 94794 94893: gap of unknown length
* 94894 96355: contig of 1462 bp in length
* 96356 96455: gap of unknown length
* 96456 97928: contig of 1473 bp in length
* 97929 98028: gap of unknown length
* 98029 99454: contig of 1426 bp in length
* 99455 99555: gap of unknown length
* 99556 101095: contig of 1541 bp in length
* 101096 101195: gap of unknown length
* 101196 102934: contig of 1739 bp in length
* 102935 103034: gap of unknown length
* 103035 104472: contig of 1438 bp in length
* 104473 104572: gap of unknown length
* 104573 106427: contig of 1855 bp in length
* 106428 106527: gap of unknown length
* 106528 108182: contig of 1655 bp in length
* 108183 108282: gap of unknown length
* 108283 110058: contig of 1776 bp in length
* 110059 110158: gap of unknown length
* 110159 111820: contig of 1662 bp in length
* 111821 111920: gap of unknown length
* 111921 113107: contig of 1187 bp in length
* 113108 113207: gap of unknown length
* 113208 114526: contig of 1319 bp in length
* 114527 114627: gap of unknown length
* 114628 116291: contig of 1665 bp in length
* 116292 116391: gap of unknown length
* 116392 117584: contig of 1193 bp in length
* 117585 117684: gap of unknown length
* 117685 118961: contig of 1277 bp in length
* 118962 119061: gap of unknown length
* 119062 120491: gap of 1429 bp in length
* 120491 120590: gap of unknown length
* 120591 121634: contig of 1044 bp in length
* 121635 121734: gap of unknown length
* 121735 123040: contig of 1306 bp in length
* 123041 123140: gap of unknown length

```

alignment\_scores:  
 Quality: 203.00 Length: 46  
 Ratio: 4.951 Gaps: 0  
 Percent Similarity: 89.130 Percent Identity: 76.087

alignment\_block:  
 US-09-877-665-4 x AC096527/rev ..  
 Align seg 1/1 to reverse of: AC096527 from: 1 to: 137600

1 HisphelysProcysArgAspLysAspLeuAlaIATryCysLeuAsnAspG1 17  
 |||||||  
 48942 CACTTCAACCTGCTGAGACAGACCTGGCGTATGCTCAAGACGCG 48933

17 yGluCysPheValIleGluThrLeuThrGlySerHisLysHisCysArgC 34  
 |||||||  
 48992 TGAATGCTTTGATGATGAGACCTTACAGATCCATAGACCTGTGCGT 48843

34 ySLySGluGlnGlyGlnGlyAlaIArgCysAspGlnPhe 46  
 :::::::::::::::  
 48842 AAGCCCTGACAGCAATAGTTAAAGTAGAGTCTCTT 48805

seg\_name: gb\_pr:AL391478

seg\_documentation\_block:

LOCUS AL391478 180625 bp DNA linear PRI 18-Apr-2001

DEFINITION Human DNA sequence from clone RP11-552113 on chromosome 10,  
 complete sequence.

ACCESSION AL391478

VERSION AL391478.14 GI:13750895

KEYWORDS HTG.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 180625)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (18-Apr-2001) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk  
 requests: clonerequests@sanger.ac.uk  
 On Apr 21, 2001 this sequence version replaced gi:13396637.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >=  
 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em: EMBL, Sw:  
 SWISSPROT, Tr: TrEMBL, Wp: WORMPEP; Information on the WORMPEP  
 database can be found at  
 http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence  
 was generated from part of bacterial clone contigs of human  
 chromosome 10, constructed by the Sanger Centre Chromosome 10  
 Mapping Group. Further information can be found at  
 http://www.sanger.ac.uk/HGP/Chr10  
 RP11-552113 is from the library RPCI-11.2 constructed by the group  
 of Pieter de Jong. For further details see  
 http://www.chori.org/Bacpac/home.htm  
 VECTOR: pBACe3.6  
 This sequence is the entire insert of clone RP11-552113 The true  
 left end of clone RP11-52203 is at 86094 in this sequence. The true  
 right end of clone RP11-319L5 is at 48826 in this sequence.  
 Location/Qualifiers

FEATURES

source

1. 180625  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="10"  
 /clone="RP11-552113"  
 /clone\_11b="RPCI-11.2"



```
repeat_region 355. .665
/note="AluX repeat: matches 1. .310 of consensus"
951. .1018
/note="34 copies 2 mer tt 70% conserved"
repeat_region 1096. .1666
/note="LIM4 repeat: matches 3695. .4279 of consensus"
1793. .2104
/note="AluY repeat: matches 1. .311 of consensus"
3064. .3099
/note="MADE1 repeat: matches 1. .36 of consensus"
3553. .3759
/note="AluSX repeat: matches 1. .283 of consensus"
5297. .5523
/note="L2 repeat: matches 1912. .2153 of consensus"
6335. .6577
/note="3 copies 81 mer 93% conserved"
7016. .7223
/note="MIR repeat: matches 56. .257 of consensus"
8275. .8558
/note="AluSX repeat: matches 10. .297 of consensus"
8559. .8735
/note="AluSg/x repeat: matches 131. .310 of consensus"
9340. .9511
/note="MIR repeat: matches 36. .219 of consensus"
9898. .10207
/note="AluY repeat: matches 1. .308 of consensus"
11032. .11317
/note="AluSX repeat: matches 1. .305 of consensus"
12667. .12793
/note="L2 repeat: matches 2612. .2750 of consensus"
13090. .13363
/note="MER4C repeat: matches 157. .1008 of consensus"
13363. .13721
/note="MER4A repeat: matches 1. .342 of consensus"
14225. .14260
/note="L2 repeat: matches 2714. .2749 of consensus"
16071. .16274
/note="LTR1 repeat: matches 1. .203 of consensus"
16289. .16420
/note="LTR28 repeat: matches 432. .558 of consensus"
16424. .16597
/note="LTR1 repeat: matches 534. .696 of consensus"
18812. .19232
/note="LIM4 repeat: matches 3410. .3826 of consensus"
19233. .19535
/note="AluSg repeat: matches 1. .303 of consensus"
19536. .20534
/note="LIM4 repeat: matches 2393. .3410 of consensus"
20542. .20777
/note="AluY repeat: matches 84. .311 of consensus"
20780. .22091
/note="LIM4 repeat: matches 1271. .2396 of consensus"
23105. .23475
/note="LTR18 repeat: matches 1. .390 of consensus"
23746. .23852
/note="LIM5A repeat: matches 6185. .6290 of consensus"
24766. .24827
/note="31 copies 2 mer tt 77% conserved"
29279. .29497
/note="AluY repeat: matches 1. .238 of consensus"
29508. .29537
/note="15 copies 2 mer aa 93% conserved"
29705. .29794
/note="L2 repeat: matches 2527. .2612 of consensus"
29997. .30450
/note="match: STS: Em:G28363"
30837. .30899
/note="MIR repeat: matches 102. .164 of consensus"
30910. .31047
/note="AluY/FRAM repeat: matches 177. .310 of consensus"
31280. .31487
/note="MER58A repeat: matches 7. .224 of consensus"
31887. .32107
repeat_region 32535. .33158
/note="AluSg/x repeat: matches 87. .312 of consensus"
/note="LIM5 repeat: matches 5333. .6002 of consensus"
34630. .34990
/note="THE1C repeat: matches 1. .371 of consensus"
36142. .36247
/note="U6 repeat: matches 1. .107 of consensus"
36248. .36509
/note="LIM2 repeat: matches 6034. .6308 of consensus"
36977. .37291
/note="LIP3 repeat: matches 5853. .6150 of consensus"
37291. .37353
/note="L1 repeat: matches 4746. .4811 of consensus"
37499. .38488
/note="LIM1 repeat: matches 4808. .5794 of consensus"
38480. .38572
/note="LIM1 repeat: matches 4746. .4836 of consensus"
38553. .39017
/note="L1 repeat: matches 4266. .4745 of consensus"
39443. .44686
/note="LIP2 repeat: matches 900. .6146 of consensus"
44682. .45457
/note="LIP2 repeat: matches 5. .776 of consensus"
45682. .45845
/note="L2 repeat: matches 2545. .2710 of consensus"
46265. .46565
/note="AluY repeat: matches 1. .297 of consensus"
47546. .47599
/note="27 copies 2 mer ac 79% conserved"
47552. .47599
/note="12 copies 4 mer acac 81% conserved"
48367. .48595
/note="L2 repeat: matches 1755. .2008 of consensus"
49387. .49709
/note="AluY repeat: matches 1. .312 of consensus"
49767. .50062
/note="AluSg repeat: matches 1. .297 of consensus"
52784. .53168
/note="LTR1A2 repeat: matches 2. .374 of consensus"
56002. .56206
/note="LTR1B repeat: matches 179. .390 of consensus"
56206. .56473
/note="LTR1B repeat: matches 1. .276 of consensus"
56935. .56965
/note="L2 repeat: matches 2672. .2702 of consensus"
complement(57033. .57459)
/note="match: STS: Em:G24088"
57451. .57492
/note="21 copies 2 mer aa 88% conserved"
57691. .57748
/note="29 copies 2 mer tg 91% conserved"
57693. .57748
/note="14 copies 4 mer ttgt 91% conserved"
57751. .58060
/note="AluY repeat: matches 1. .294 of consensus"
58815. .59127
/note="AluY repeat: matches 1. .311 of consensus"
63084. .63401
/note="AluSX repeat: matches 1. .310 of consensus"
63416. .63529
/note="MIR repeat: matches 110. .240 of consensus"
complement(64102. .64292)
/note="match: STS: Em:G03280"
64682. .64970
/note="AluY repeat: matches 1. .288 of consensus"
65253. .66003
/note="LTR1 repeat: matches 1. .785 of consensus"
66317. .66372
/note="14 copies 4 mer acac 85% conserved"
66979. .67390
/note="MIR repeat: matches 2. .418 of consensus"
67989. .68264
/note="AluSX repeat: matches 1. .276 of consensus"
```

Repeat\_region 69790..69889

## alignment\_scores:

Quality: 195.00 Length: 33  
Ratio: 5.909 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-877-665-4 x AL391478 ..

Align seg 1/1 to: AL391478 from: 1 to: 180625

1 HisphelysProcysargaspIysaspIleuAlaIatYrCysIleuasnspgl 17  
|||||  
54142 CACTCAAAACCTGCGGACAGACAGACCTGCTACTGCTCATATGAG 54191  
|||||

17 yglucysphevalIleGlutThrleuthrIglYserHIScysarg 33  
|||||

54192 CGAGTGTGTGATCGAAACCTGACCGATCCCATAAACACTGTGCG 54240  
|||||

seq\_name: gb\_ov:AF076618

## seq\_documentation\_block:

LOCUS AF076618 3105 bp mRNA linear VRT 18-JUL-1998  
DEFINITION Xenopus laevis neuregulin alpha-1 mRNA, complete cds.  
ACCESSION AF076618  
VERSION AF076618.1 GI:3328216  
KEYWORDS  
SOURCE African clawed frog.  
ORGANISM Xenopus laevis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;  
Xenopodinae; Xenopus.  
REFERENCE 1 (bases 1 to 3105)  
Yang, J.F., Zhou, H., Pun, S., Ip, N.Y., Peng, H.B. and Tsai, K.W.  
Cloning of cDNAs encoding xenopus neuregulin: expression in  
myotomal muscle during embryo development  
JOURNAL Brain Res. Mol. Brain Res. 58 (1-2), 59-73 (1998)  
MEDLINE 98352126  
REFERENCE 2 (bases 1 to 3105)  
Yang, J.F., Zhou, H., Pun, S., Ip, N.Y., Peng, H.B. and Tsai, K.W.  
Direct Submission  
Submitted (03-JUL-1998) Department of Biology and Biotechnology  
Institute, The Hong Kong University of Science and Technology,  
Clear Water Bay Road, Hong Kong, China  
JOURNAL Location/Qualifiers  
1. 3105  
/organism="Xenopus laevis"  
/db\_xref="taxon:8355"  
161. 2194  
/function="acetylcholine receptor-inducing activity"  
/note="SP43"  
/codon\_start=1  
/product="neuregulin alpha-1"  
/protein\_id="AAC26804.1"  
/db\_xref="GI:3328217"

## FEATURES

source

CDS

BASE COUNT

1098 a 624 c 635 g 748 t

alignment\_scores:

Quality: 126.50 Length: 46  
Ratio: 3.614 Gaps: 1  
Percent Similarity: 76.087 Percent Identity: 43.478

## alignment\_block:

US-09-877-665-4 x AF076618 ..

Align seg 1/1 to: AF076618 from: 1 to: 3105

1 HisphelysProcysargaspIysaspIleuAlaIatYrCysIleuasnspgl 17  
|||||  
722 CACCTTATTAAGTTCACACACAGAAAGACCTACTGTCTCATGAGG 771  
|||||

17 yglucysphevalIleGlutThrleuthrIglYserHIScysarg 33  
|||||

772 AGAGTGTATGTACTCAGCGATATACAGACAGACAGACGTTTCATGCA 821  
|||||

33 rgcysIysgluIglYrGlnGlyValArgCysasnpgln 45  
|||||

822 AGTGCAAACCTGATCTACTGAGACAGATGACACTGAA 859  
|||||

seq\_name: gb\_pat:AR100663

## seq\_documentation\_block:

LOCUS AR100663 156 bp DNA linear PAT 14-FEB-2001  
DEFINITION Sequence 2 from patent US 6080845.  
ACCESSION AR100663  
VERSION AR100663.1 GI:12811111  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 156)  
Carnahan, J.F.  
TITLE Monoclonal antibody against utricular epithelium  
JOURNAL Patent: US 6080845-A 2 27-JUN-2000;  
FEATURES  
1. 156  
Location/Qualifiers  
source  
/organism="unknown"

BASE COUNT 47 a 31 c 42 g 36 t  
ORIGIN

## alignment\_scores:

Quality: 116.50 Length: 48  
Ratio: 3.149 Gaps: 1  
Percent Similarity: 77.083 Percent Identity: 33.333

## alignment\_block:

US-09-877-665-4 x AR100663 ..

Align seg 1/1 to: AR100663 from: 1 to: 156

1 HisphelysProcysargaspIysaspIleuAlaIatYrCysIleuasnspgl 17  
|||||  
4 CATCTTAAATAAGTTCGCGAAGAGAAACCTTCTGTGATGAGAG 53  
|||||

17 yglucysphevalIleGlutThrleuthrIglYserHIScysarg 33  
|||||

54 GGAATGCTTATGCTGAAAGACCTTCAACCCCTCGAATACTTGCA 103  
|||||

33 rgcysIysgluIglYrGlnGlyValArgCysasnpgln 47  
|||||

104 AGTGCAAACCTGATCTACTGAGACAGATGTCAAACCTACGTA 147  
|||||

seq\_name: gb\_pat:AR022501

## seq\_documentation\_block:

LOCUS AR022501 192 bp DNA linear PAT 05-DEC-1998  
DEFINITION Sequence 155 from patent US 5792849.  
ACCESSION AR022501  
VERSION AR022501.1 GI:3976563  
KEYWORDS



LOCUS	AR143352	192 bp	DNA	linear	PAT 08-AUG-2001
DEFINITION	Sequence 155 from patent US 6204241.				
ACCESSION	AR143352				
VERSION	AR143352.1	GI:15104638			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 192)				
TITLE	Goodearl,A.David, Stroobant,P., Minghetti,L., Waterfield,M., Marchionni,M., Chen,M.Su and Hiles,I.				
JOURNAL	Method for treating nervous system pathophysologies using glial growth factors				
FEATURES	Patent: US 6204241-A 155 20-MAR-2001;				
source	Location/Qualifiers				
BASE COUNT	1..192				
ORIGIN	/organism="unknown"				
	61 a	39 c	50 g	42 t	
alignment_scores:	Quality: 113.50	Length: 46			
	Ratio: 3.338	Gaps: 1			
Percent Similarity:	73.913	Percent Identity:	34.783		
alignment_block:	US-09-877-665-4 x AR143352 ..				
Align seg 1/1 to:	AR143352	from: 1	to: 192		
1	HisPheIysProCysArgAspIysAspLeuAlaItyrCysLeuAsnAspI 17				
	:     :   :				
4	CACTCTGCAAGTGTGCAGAGAAGAAACTTCTGTGTGAATGAGG 53				
	: :				
17	yglucysPheValIleGluThrIleuThrGlySerHisIstLysHis..Cysa 33				
	: :				
54	CGAGTCTCTCATGTGTAAGACCTTTCAAATCCCTCAAAATACTTGTGCA 103				
	: :				
33	rgCysIysGluIuItyrGInglyValArgCysAspIin 45				
	: :				
104	AGTGCCAACCTGTGATTCACGTGAGCGAGCATGTACTGTAG 141				
seq_name:	gb_pat:AR151961				
seq_documentation_block:					
LOCUS	AR151961	192 bp	DNA	linear	PAT 08-AUG-2001
DEFINITION	Sequence 155 from patent US 6232286.				
ACCESSION	AR151961				
VERSION	AR151961.1	GI:15118011			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 192)				
TITLE	Goodearl,A., Stroobant,P., Minghetti,L., Waterfield,M., Marchionni,M., Chen,M.Su and Hiles,I.				
JOURNAL	Methods of stimulating mitogenesis in glial cells using glial mtogenic factors				
FEATURES	Patent: US 6232286-A 155 15-MAY-2001;				
source	Location/Qualifiers				
BASE COUNT	1..192				
ORIGIN	/organism="unknown"				
	61 a	39 c	50 g	42 t	
Alignment_scores:	Quality: 113.50	Length: 46			
	Ratio: 3.338	Gaps: 1			
Percent Similarity:	73.913	Percent Identity:	34.783		





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 08:42:13 ; Search time 13.02 Seconds

(without alignments)  
88.172 Million cell updates/sec

Title: US-09-877-665-4

Sequence: 1 HFKRCRDLAYCLNDGECF.....SHKRCRKEGYGVRCDOFL 47

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 segs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA:\*

- 1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*\*
- 2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*\*
- 3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*\*
- 4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*\*
- 5: /cgn2\_6/ptodata/2/1aa/PCRTGS.COMB.pep:\*\*
- 6: /cgn2\_6/ptodata/2/1aa/Backfiles1.pep:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	277	100.0	47	3	US-08-899-437-4
2	277	100.0	47	3	US-08-899-437-8
3	277	100.0	47	4	US-09-126-121-4
4	277	100.0	47	4	US-09-126-121-8
5	277	100.0	360	3	US-08-899-437-7
6	277	100.0	360	4	US-09-126-121-7
7	277	100.0	362	3	US-08-899-437-3
8	277	100.0	362	4	US-09-126-121-3
9	277	100.0	696	3	US-08-899-437-23
10	277	100.0	696	4	US-09-126-121-23
11	277	100.0	713	3	US-08-899-437-2
12	277	100.0	713	4	US-09-126-121-2
13	277	100.0	720	3	US-08-899-437-6
14	277	100.0	720	4	US-09-126-121-6
15	116.5	42.1	52	1	US-08-417-640A-1
16	116.5	42.1	52	1	US-08-760-815-1
17	116.5	42.1	52	2	US-08-761-038-1
18	116.5	42.1	52	3	US-09-238-182-1
19	113.5	41.0	49	3	US-08-899-437-14
20	113.5	41.0	49	4	US-09-126-121-14
21	113.5	41.0	50	3	US-08-753-007A-12
22	113.5	41.0	50	4	US-09-398-496-12
23	113.5	41.0	52	1	US-08-417-640A-3
24	113.5	41.0	52	1	US-08-760-815-3
25	113.5	41.0	52	2	US-08-761-038-3
26	113.5	41.0	54	1	US-08-179-481-111
27	113.5	41.0	63	3	US-08-341-018-62

28	113.5	41.0	63	4	US-08-470-335-221	Sequence 221, App
29	113.5	41.0	63	4	US-08-470-339-221	Sequence 221, App
30	113.5	41.0	66	1	US-07-847-743B-10	Sequence 10, App
31	113.5	41.0	66	1	US-08-456-201-10	Sequence 10, App
32	113.5	41.0	66	2	US-08-456-241-10	Sequence 10, App
33	113.5	41.0	66	4	US-09-020-880-2	Sequence 2, App
34	113.5	41.0	66	5	PCT-US92-04295A-10	Sequence 10, App
35	113.5	41.0	83	3	US-08-341-018-70	Sequence 225, App
36	113.5	41.0	83	4	US-08-470-335-225	Sequence 225, App
37	113.5	41.0	88	3	US-08-470-339-225	Sequence 68, App
38	113.5	41.0	88	4	US-08-470-335-224	Sequence 224, App
39	113.5	41.0	88	4	US-08-470-339-224	Sequence 224, App
40	113.5	41.0	88	4	US-08-470-339-224	Sequence 70, App
41	113.5	41.0	95	1	US-07-847-743B-14	Sequence 14, App
42	113.5	41.0	95	1	US-08-456-201-14	Sequence 12, App
43	113.5	41.0	95	2	US-08-330-161-12	Sequence 14, App
44	113.5	41.0	95	2	US-08-456-241-14	Sequence 12, App
45	113.5	41.0	95	2	US-08-440-401-12	Sequence 12, App

## ALIGNMENTS

*Some protein*

RESULT 1  
US-08-899-437-4  
Sequence 4, Application US/08899437  
Patent No. 6121415  
GENERAL INFORMATION:  
APPLICANT: GIGOSKI, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESSES:  
ADDRESS: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/899,437  
FILING DATE: 24-Jul-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Conley, Delidre L.  
REGISTRATION NUMBER: 36,487  
REFERENCE/DOCKET NUMBER: P1084R1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-2066  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 47 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
FEATURE:  
NAME/KEY: NRG3 EGF-like domain/amino acid seq.  
LOCATION: 1-47  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
US-08-899-437-4

Query Match 100.0%; Score 277; DB 3; Length 47;  
Best Local Similarity 100.0%; Pred. No. 1.3e-26;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 1 HFKRCRDLAYCLNDGECFVETLTGSHKRCRKEGYGVRCDOFL 47

DB 1 HFKPCRDILAYCLNDEGCFYIEITLTSKHKCRCKEGYGVRCDOFL 47

RESULT 2

US-08-899-437-8 Application US/08899437

Sequence 8 6121415

Patent No. 6121415

GENERAL INFORMATION:

APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao

TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/899,437

FILING DATE: 24-Jul-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Conley, Delidre L.

REGISTRATION NUMBER: 36,487

REFERENCE/DOCKET NUMBER: P1084R1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/952-2066

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 47 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

FEATURE:

NAME/KEY: NR3 EGF-like domain/amino acid seq.

LOCATION: 1-47

IDENTIFICATION METHOD:

OTHER INFORMATION:

US-08-899-437-8

Query Match 100.0%; Score 277; DB 3; Length 47;

Best Local Similarity 100.0%; Pred. No. 1.3e-26;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 HFKPCRDILAYCLNDEGCFYIEITLTSKHKCRCKEGYGVRCDOFL 47

RESULT 3

US-09-126-121-4

Sequence 4 Application US/09126121

Patent No. 6252051

GENERAL INFORMATION:

APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao

TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/126,121

FILING DATE: 30-Jul-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Conley, Delidre L.

REGISTRATION NUMBER: 36,487

REFERENCE/DOCKET NUMBER: P1084R1D1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/952-2066

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 47 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

FEATURE:

NAME/KEY: NR3 EGF-like domain/amino acid seq.

LOCATION: 1-47

IDENTIFICATION METHOD:

OTHER INFORMATION:

US-09-126-121-4

Query Match 100.0%; Score 277; DB 4; Length 47;

Best Local Similarity 100.0%; Pred. No. 1.3e-26;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 HFKPCRDILAYCLNDEGCFYIEITLTSKHKCRCKEGYGVRCDOFL 47

RESULT 4

US-09-126-121-8

Sequence 8 Application US/09126121

Patent No. 6252051

GENERAL INFORMATION:

APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao

TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/126,121

FILING DATE: 30-Jul-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Conley, Delidre L.

REGISTRATION NUMBER: 36,487

REFERENCE/DOCKET NUMBER: P1084R1D1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/952-2066

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:



LENGTH: 47 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
FEATURE:  
NAME/KEY: NR3 EGF-like domain/amino acid seq.  
LOCATION: 1-47  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
US-09-126-121-8

Query Match 100.0%; Score 277; DB 4; Length 47;  
Best Local Similarity 100.0%; Pred. No. 1.3e-26;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFKPCRDKLAVCLNDGCEFYETLTGSHKRCRCKEGYGVRCDFL 47  
DB 1 HFKPCRDKLAVCLNDGCEFYETLTGSHKRCRCKEGYGVRCDFL 47

RESULT 5  
US-08-899-437-7  
Sequence 7, Application US/08899437  
Patent No. 6121415  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
TITLE OF INVENTION: Ligands and Uses Therefor  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/899,437  
FILING DATE: 24-Jul-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Conley, Delaire L.  
REGISTRATION NUMBER: 36,487  
REFERENCE/DOCKET NUMBER: P1084R1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-2066  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 360 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
FEATURE:  
NAME/KEY: hNRG3 extracellular domain/Amino AcidSeq  
LOCATION: 1-360  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
US-08-899-437-7

Query Match 100.0%; Score 277; DB 3; Length 360;  
Best Local Similarity 100.0%; Pred. No. 1e-25;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFKPCRDKLAVCLNDGCEFYETLTGSHKRCRCKEGYGVRCDFL 47  
DB 286 HFKPCRDKLAVCLNDGCEFYETLTGSHKRCRCKEGYGVRCDFL 332

RESULT 6  
US-09-126-121-7  
Sequence 7, Application US/09126121  
Patent No. 6252051  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
TITLE OF INVENTION: Ligands and Uses Therefor  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/126,121  
FILING DATE: 30-Jul-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Conley, Delaire L.  
REGISTRATION NUMBER: 36,487  
REFERENCE/DOCKET NUMBER: P1084R1D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-2066  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 360 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
FEATURE:  
NAME/KEY: hNRG3 extracellular domain/Amino AcidSeq  
LOCATION: 1-360  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
US-09-126-121-7

Query Match 100.0%; Score 277; DB 4; Length 360;  
Best Local Similarity 100.0%; Pred. No. 1e-25;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFKPCRDKLAVCLNDGCEFYETLTGSHKRCRCKEGYGVRCDFL 47  
DB 286 HFKPCRDKLAVCLNDGCEFYETLTGSHKRCRCKEGYGVRCDFL 332

RESULT 7  
US-08-899-437-3  
Sequence 3, Application US/08899437  
Patent No. 6121415  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
TITLE OF INVENTION: Ligands and Uses Therefor  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatln (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/899,437  
FILING DATE: 24-Jul-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Conley, Delidre L.  
REGISTRATION NUMBER: 36,487  
REFERENCE/DOCKET NUMBER: P1084R1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-2066  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 362 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
FEATURE:  
NAME/KEY: MNRG3 extracellular domain amino acid seq  
LOCATION: 1-362  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
US-08-899-437-3

Query Match 100.0%; Score 277; DB 3; Length 362;  
Best Local Similarity 100.0%; Pred. No. 1e-25;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HFKPCRDIAVCLNDGECFVETLTGSHKRCRCEGYGVRCDFL 47  
DB 288 HFKPCRDIAVCLNDGECFVETLTGSHKRCRCEGYGVRCDFL 334

RESULT 8  
US-09-126-121-3  
Sequence 3, Application US/09126121  
Patent No. 6252051  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
TITLE OF INVENTION: Ligands and Uses Therefor  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatln (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/126,121  
FILING DATE: 30-Jul-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Conley, Delidre L.  
REGISTRATION NUMBER: 36,487  
REFERENCE/DOCKET NUMBER: P1084R1D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-2066  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 362 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear

FEATURE:  
NAME/KEY: MNRG3 extracellular domain amino acid seq  
LOCATION: 1-362  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
US-09-126-121-3

Query Match 100.0%; Score 277; DB 4; Length 362;  
Best Local Similarity 100.0%; Pred. No. 1e-25;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HFKPCRDIAVCLNDGECFVETLTGSHKRCRCEGYGVRCDFL 47  
DB 288 HFKPCRDIAVCLNDGECFVETLTGSHKRCRCEGYGVRCDFL 334

RESULT 9  
US-08-899-437-23  
Sequence 23, Application US/08899437  
Patent No. 6121415  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
TITLE OF INVENTION: Ligands and Uses Therefor  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatln (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/899,437  
FILING DATE: 24-Jul-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Conley, Delidre L.  
REGISTRATION NUMBER: 36,487  
REFERENCE/DOCKET NUMBER: P1084R1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-2066  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 696 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
FEATURE:  
NAME/KEY: Human MNRG3B2  
LOCATION: 1-696  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
US-08-899-437-23

Query Match 100.0%; Score 277; DB 3; Length 696;  
Best Local Similarity 100.0%; Pred. No. 2e-25;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HFKPCRDIAVCLNDGECFVETLTGSHKRCRCEGYGVRCDFL 47  
DB 286 HFKPCRDIAVCLNDGECFVETLTGSHKRCRCEGYGVRCDFL 332

RESULT 10  
US-09-126-121-23

Sequence 23, Application US/09126121  
Patent No. 6252051  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/126,121  
FILING DATE: 30-Jul-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Conley, Delidre L.  
REGISTRATION NUMBER: 36,487  
REFERENCE/DOCKET NUMBER: P1084R1D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-2066  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 696 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
FEATURE:  
NAME/KEY: Human NRG3B2  
LOCATION: 1-696  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
US-09-126-121-23

Query Match 100.0%; Score 277; DB 4; Length 696;  
Best Local Similarity 100.0%; Pred. No. 2e-25;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HFKPCRDKLATCLNDGECFVETLTGSHKCRCKRGYGVGRDQFL 47  
|||||  
DB 286 HFKPCRDKLATCLNDGECFVETLTGSHKCRCKRGYGVGRDQFL 332

RESULT 11  
US-08-899-437-2  
Sequence 2, Application US/08899437  
Patent No. 6121415  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/899,437  
FILING DATE: 24-Jul-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Conley, Delidre L.  
REGISTRATION NUMBER: 36,487  
REFERENCE/DOCKET NUMBER: P1084R1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-2066  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 713 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
FEATURE:  
NAME/KEY: Mouse NRG3 (mNRG3)/amino acid seq.  
LOCATION: 1-713  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
US-08-899-437-2

Query Match 100.0%; Score 277; DB 3; Length 713;  
Best Local Similarity 100.0%; Pred. No. 2.1e-25;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HFKPCRDKLATCLNDGECFVETLTGSHKCRCKRGYGVGRDQFL 47  
|||||  
DB 288 HFKPCRDKLATCLNDGECFVETLTGSHKCRCKRGYGVGRDQFL 334

RESULT 12  
US-09-126-121-2  
Sequence 2, Application US/09126121  
Patent No. 6252051  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/126,121  
FILING DATE: 30-Jul-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Conley, Delidre L.  
REGISTRATION NUMBER: 36,487  
REFERENCE/DOCKET NUMBER: P1084R1D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-2066  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 713 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
FEATURE:  
NAME/KEY: Mouse NRG3 (mNRG3)/amino acid seq.  
LOCATION: 1-713

IDENTIFICATION METHOD:  
OTHER INFORMATION:  
US-09-126-121-2

Query Match 100.0%; Score 277; DB 4; Length 713;  
Best Local Similarity 100.0%; Pred. No. 2.1e-25;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HFKPCRDRLAYCLNDECFVIEITLGSKHKCRCKEGYGVRCDOFL 47  
DB 288 HFKPCRDRLAYCLNDECFVIEITLGSKHKCRCKEGYGVRCDOFL 334

RESULT 13  
US-08-899-437-6  
Sequence 6, Application US/08899437  
Patent No. 6121415

GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESS: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/899,437  
FILING DATE: 24-Jul-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Conley, Delidre L.  
REGISTRATION NUMBER: 36,487  
REFERENCE/DOCKET NUMBER: P1084R1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-2066  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 720 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear

FEATURE:  
NAME/KEY: hNR3B1 amino acid sequence  
LOCATION: 1-720  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
US-08-899-437-6

Query Match 100.0%; Score 277; DB 3; Length 720;  
Best Local Similarity 100.0%; Pred. No. 2.1e-25;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HFKPCRDRLAYCLNDECFVIEITLGSKHKCRCKEGYGVRCDOFL 47  
DB 286 HFKPCRDRLAYCLNDECFVIEITLGSKHKCRCKEGYGVRCDOFL 332

RESULT 14  
US-09-126-121-6  
Sequence 6, Application US/09126121  
Patent No. 6252051  
GENERAL INFORMATION:  
US-09-126-121-6

IDENTIFICATION METHOD:  
OTHER INFORMATION:  
US-09-126-121-6

APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESS: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/126,121  
FILING DATE: 30-Jul-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Conley, Delidre L.  
REGISTRATION NUMBER: 36,487  
REFERENCE/DOCKET NUMBER: P1084R1D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-2066  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 720 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear

FEATURE:  
NAME/KEY: hNR3B1 amino acid sequence  
LOCATION: 1-720  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
US-09-126-121-6

Query Match 100.0%; Score 277; DB 4; Length 720;  
Best Local Similarity 100.0%; Pred. No. 2.1e-25;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HFKPCRDRLAYCLNDECFVIEITLGSKHKCRCKEGYGVRCDOFL 47  
DB 286 HFKPCRDRLAYCLNDECFVIEITLGSKHKCRCKEGYGVRCDOFL 332

RESULT 15  
US-08-417-640A-1  
Sequence 1, Application US/08417640A  
Patent No. 5670342  
GENERAL INFORMATION:  
APPLICANT: Carnahan, Josette F.  
APPLICANT: Hara, Shinichi  
APPLICANT: Lu, Hsiang S.  
APPLICANT: Mayer, John P.  
APPLICANT: Yoshinaga, Steven K.  
TITLE OF INVENTION: NDF Peptides  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESS: Amgen Inc.  
STREET: 1840 Denavilland Drive  
CITY: Thousand Oaks  
STATE: California  
COUNTRY: USA  
ZIP: 91320

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 08:49:33 Search time 25.6 Seconds  
(without alignments)  
317,608 Million cell updates/sec

Title: US-09-877-665-4  
Perfect score: 277  
Sequence: 1 HFKPCRDIAVCINDGECF.....SHKRCCKEKGVCACDPL 47

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

1: SP archaea:\*  
2: SP bacteria:\*  
3: SP fungi:\*  
4: SP human:\*  
5: SP\_invertebrate:\*  
6: SP\_mammal:\*  
7: SP\_mhc:\*  
8: SP\_organelle:\*  
9: SP\_phage:\*  
10: SP\_plant:\*  
11: SP\_rodent:\*  
12: SP\_virus:\*  
13: SP Vertebrate:\*  
14: SP\_unclassified:\*  
15: SP\_virus:\*  
16: SP\_bacterioplasmid:\*  
17: SP\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	113.5	41.0	461	11	035947 mesocricetu
2	111.5	40.3	298	11	09ESa9 rattus norv
3	111.5	40.3	695	11	09ESB0 rattus norv
4	104.5	37.7	241	6	007112 bos taurus
5	104.5	37.7	296	4	0961B3
6	102.5	37.0	111	11	09ESa8 rattus norv
7	102.5	37.0	136	11	09ESa7 rattus norv
8	102.5	37.0	156	11	09ESa6 rattus norv
9	102.5	37.0	317	11	09ESa3 rattus norv
10	102.5	37.0	323	11	09ESa2 rattus norv
11	102.5	37.0	342	11	09ESa1 rattus norv
12	102.5	37.0	700	11	09ESB1 rattus norv
13	102.5	37.0	782	11	09ESa5 rattus norv
14	99	35.7	89	12	091M20
15	92.5	33.4	2192	5	001768
16	91.5	33.0	1213	13	090Y54

17	89	32.1	162	11	0920L5	0920L5 rattus norv
18	89	32.1	1241	4	09UKK5	09UKK5 homo sapien
19	89	32.1	1241	4	09UKD4	09UKD4 homo sapien
20	89	32.1	1241	4	09B2V3	09B2V3 homo sapien
21	88	31.8	1239	11	P70628	P70628 rattus norv
22	87	31.4	217	5	09VVT6	09VVT6 diosophila
23	87	31.4	597	11	035727	035727 mus muscullu
24	85.5	30.9	1193	13	090819	090819 gallus gall
25	85.5	30.9	1214	13	090YD2	090YD2 xenopus lae
26	85.5	30.9	1218	4	014902	014902 homo sapien
27	85.5	30.9	1218	4	015122	015122 homo sapien
28	85.5	30.9	1218	4	015816	015816 homo sapien
29	85.5	30.9	1227	4	P78504	P78504 homo sapien
30	84.5	30.5	177	11	09JTM4	09JTM4 rattus norv
31	84.5	30.5	1218	11	09QXX0	09QXX0 mus muscullu
32	84.5	30.5	1219	11	063722	063722 rattus norv
33	83.5	30.1	127	12	0909F3	0909F3 ectomella
34	83	30.0	162	11	061521	061521 mus muscullu
35	83	30.0	861	11	09QW58	09QW58 mus sp. mot
36	83	30.0	862	11	099JC2	099JC2 mus muscullu
37	82.5	29.8	445	5	09W3W5	09W3W5 diosophila
38	82.5	29.8	456	5	0961F3	0961F3 diosophila
39	82	29.6	264	5	020559	020559 caenorhabd1
40	81	29.2	78	12	091T36	091T36 lumpy skin
41	80.5	29.1	125	12	041504	041504 compox viru
42	80.5	29.1	178	4	096F48	096F48 homo sapien
43	80.5	29.1	260	4	095898	095898 homo sapien
44	80.5	29.1	530	4	096CM9	096CM9 homo sapien
45	80.5	29.1	1242	13	090Y57	090Y57 brachydantio

#### ALIGNMENTS

RESULT 1  
ID 035947 PRELIMINARY; PRT; 461 AA.  
AC 035947;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE PRO-NEUREGULIN-1, ISOFORM ALPHA 2B PRECURSOR.  
GN NG1 OR NDF.  
OS Mesocricetus auratus (Golden hamster).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
OC Mesocricetus.  
OX NCBI\_Taxid=10036;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM ALPHA2B), AND SEQUENCE OF 64-81.  
RC TISSUE=EMBRYO;  
RX MEDLINE-98196966; PubMed-9537646;  
RA Velasco J.A., Feljoo E., Avila M.A., Notario V.;  
RT Secretion of new differentiation factor-like polypeptides by cph-  
RT transformed fibroblasts: cloning and characterization of Syrian  
RT hamster neurogenin cDNAs.  
RT Mol. Carcinog. 21:156-163(1998).  
RL  
-1- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE  
RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,  
RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND  
ACTIVATION OF THE ERBB RECEPTORS. MAY PLAY AN IMPORTANT ROLE IN  
PROVIDING GROWTH ADVANTAGE IN NEOPLASTIC CELLS.  
-1- SUBUNIT: THE CYTOPLASMIC DOMAIN INTERACTS WITH THE LIM DOMAIN  
REGION OF LIM1 (BY SIMILARITY).  
-1- SUBCELLULAR LOCATION: EXISTS AS TYPE I MEMBRANE PROTEIN AND AS A  
PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE MEMBRANE-  
BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).  
-1- ALTERNATIVE PRODUCTS: DIFFERENT ISOFORMS ARE PRODUCED BY  
ALTERNATIVE SPLICING. THE SEQUENCE SHOWN IS THAT OF ISOFORM  
ALPHA2B/CLONE PM3.  
-1- TISSUE SPECIFICITY: EXPRESSED AT HIGHER LEVEL AFTER NEOPLASMIC  
TRANSFORMATION OF CELLS.  
-1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION

CC		OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE
CC		PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN
CC		DIMERIZATION (BY SIMILARITY).
CC	-1-	DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-L
CC		DOMAIN.
CC	-1-	PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE
CC		EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR
CC		(BY SIMILARITY).
CC	-1-	PTM: EXTENSIVE GLYCOSYLATION PRECEDES PROTEOLYTIC CLEAVAGE (BY
CC		SIMILARITY).
CC	-1-	SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC	-1-	SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC		-1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
DR	EMBL; U96612;	AAB71812.1; "
DR	HSSP; Q12784;	IHR.
DR	Interpro: IPR0000561;	EGF-like.
DR	Interpro: IPR003598;	Ig_c2.
DR	Interpro: IPR003006;	Ig_MHC.
DR	Interpro: IPR002154;	Neuregulin.
DR	Pfam; PF00008;	Egf. 1.
DR	Pfam; PF00047;	Ig. 1.
DR	Pfam; PFO2158;	Neuregulin; 1.
DR	PRINTS; PR01089;	NEUREGULIN.
DR	SMART; SM00181;	EGF. 1
DR	SMART; SM00408;	IGC2. 1.
DR	PROSITE; PS00022;	EGF-1; UNKNOWN_1.
DR	Growth factor; EGf-1like domain;	Immunoglobulin domain; Glycoprotein,
KW	Transmembrane; Alternative splicing.	
KW	PROPEP	1 13
FT	CHAIN	14 461
FT	CHAIN	14 241.
FT	DOMAIN	14 242
FT	TRANSLEM	243 265
FT	DOMAIN	266 461
FT	DOMAIN	50 119
FT	DOMAIN	165 177
FT	DOMAIN	178 232
FT	DISULEFD	57 112
FT	DISULEFD	182 196
FT	DISULEFD	190 210
FT	DISULEFD	212 221
FT	CARBOHYD	73 73
FT	CARBOHYD	120 120
FT	CARBOHYD	126 126
FT	CARBOHYD	164 164
SO	SEQUENCE	461 AA; 50890 MW; 935CG9560F7148336 CRC4;

	Query Match	41.0%;	Score 113.5;	DB 11;	Length 461;
	Best Local Similarity	34.8%;	Pred. No. 1e-07;		
	Matches 16;	Conservative 14;	Mismatches 15;	Indels 1;	Gaps 1
Oy	1 HEKPCRDLDLAYCLNDGCECFYETLTGSHKH-CRCKEGYQVGRCDQ 45     : : :     :::   : :   :     :				
Dd	178 HLVKCAEKEKCTPCVNGGECPFWKDLSNPRIYCKRCQPETTGARCTE 223				
RESULT	2				
Q9ESA9					
ID Q9ESA9	PRELIMINARY;	PRF;	298 AA.		
AC Q9ESA9;					
DT 01-MAR-2001 (TReMBLrel. 16; Created)					
DT 01-MAR-2001 (TReMBLrel. 16; Last sequence update)					
DT 01-DEC-2001 (TReMBLrel. 19; Last annotation update)					
DE SMPF NEUREGUCLIN ALPHA 2B (FRAGMENT).					
GN NRGL.					
OS Rattus norvegicus (Rat).					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Rattus.					
OX NCBI_TaxID=10116;					
RN [1]					
RP SEQUENCE FROM N.A.					

RC	:STRAIN-BDIX;
RA	Carroll S.L., Anderson K.D., Frohnert P.W.;
RT	"Structural and Functional Diversity of SMDR Neuregulin Splice
RL	Variants Expressed in the Adult Rat Nervous System.";
DR	Submitted (OCT-1999) to the EMBL/genbank/DBDB databases.
DR	EMBL; AF194440; AAGS8429.1; -
DR	HSSP; O12784; IHRF.
DR	InterPro; IPRO00561; EGF-like.
DR	InterPro; IPRO02154; Neuregulin.
DR	InterPro; IPRO02114; PTS_HPR_ser.
DR	Pfam; PF00008; EGF_1.
DR	Pfam; PF02158; Neuregulin; 1.
DR	PRINTS; PRO1089; NEUREGULIN.
DR	SMART; SM00181; EGF_1.
DR	SMART; SM00001; EGF_like; 1.
DR	PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR	PROSITE; PSS00589; PTS_HPR_SER; UNKNOWN_1.
FT	NON_TER
FT	NON_TER
SO	SEQUENCE

[illegible]

RESULT	3	
ID	09ESB0	PRELIMINARY; PRT; 695 AA.
AC	09ESB0.	
DT	01-MAR-2001 (TREMBLrel. 16, Created)	
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)	
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)	
DE	SMDF NEUREGULIN ALPHA 2A.	
GN	NRG1.	
OS	Rattus norvegicus (Rat).	
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
OX	NCBI_TaxID=10116;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN-BDIX;	
RA	Carroll S.L., Anderson K.D., Frohner P.W.;	
RT	"Structural and Functional Diversity of SMGF Neuregulin Splice Variants Expressed in the Adult Rat Nervous System.";	
RL	Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.	
DR	EMBL; AF194439; AAC628428.1; -	
DR	HSSP; Q12784; IHRE.	
DR	InterPro: IPR000561; EGF_1like.	
DR	InterPro: IPR002154; Neuregulin.	
DR	InterPro: IPR002114; PRT_HPR_ser.	
DR	Pfam; PF00008; EGF_1.	
DR	Pfam; PF02158; Neuregulin; 1.	
DR	PRINTS; PRO1089; NEUREGULIN.	
DR	SMART; SM00181; EGF_1.	
DR	SMART; SM00001; EGF_1like; 1.	
DR	PROSITE: PS00023; EGF_1; UNKNOWN.1.	
DR	PROSITE: PS00589; PRT_HPR_SER; UNKNOWN.1.	
SO	SEQUENCE 695 AA; 5277F2CBA2FB6878 CRC64;	

Query Match 40.3% Score 11.5 DB 11; Length 69;  
Best Local Similarity 34.8%; Pred. 3.1e-07;  
Matches 16; Conservative 13; Mismatches 16; Indels 1;  
Gaps 1;

QY 1 HFKPRDKDLAYCLNDGECFVLETITGSHKH-CRCKEGYQGVRCDQ 45  
| : : : : :  
Db 234 HLLCAEKEKFTCVNGSGCFYVKDLSNFSRYLCKCPQFTFGARCE 279  
| : : : : :  
| : : : : :

RESULT	ID	PRELIMINARY:	PRT:	241 AA.
4	007112			
AC	007112			
DT	01-JAN-1998 (TREMBLrel). 05, Created)			
DT	01-JAN-1998 (TREMBLrel). 05, last sequence update)			
DT	01-JUN-2001 (TREMBLrel). 17, last annotation update)			
DE	NEUREGULIN-1, GLIAL GROWTH FACTOR 5 ISOFORM PRECURSOR (GGFBPP5).			
CN	NGFI OR GGF.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovine; Bos.			
OX	NCBI_TaxId=9913;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RC	TISSUE=POSTERIOR PITUITARY;			
RX	MEDLINE=93205115; PubMed=8096067;			
RA	Marcloni M.A., Gooden A.D.J., Chen M.S., Bermingham-McDonogh O.,			
RA	Kirk C., Hendricks M., Danely F., Misumi D., Sudhalter J.,			
RA	Kobayashi K., Wroblewski D., Lynch C., Baldassarre M., Hiles I.,			
RA	Davis J.B., Hsuan J.J., Totty N.C., Otsu M., McBurney R.N.,			
RT	Waterfield M.D., Stroobant P., Gwynne D.;			
RT	"Glial growth factors are alternatively spliced erbB2 ligands			
RT	expressed in the nervous system." ,			
RL	Nature 362:312-318(1993).			
CC	- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE			
CC	RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,			
CC	RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND			
CC	ACTIVATION OF THE ERBB RECEPTORS (BY SIMILARITY).			
CC	- SUBCELLULAR LOCATION: SECRETED (PROBABLE).			
CC	- TISSUE SPECIFICITY: EXPRESSED IN NERVOUS SYSTEM: SPINAL CORD AND			
CC	BRAIN			
CC	- ALTERNATIVE PRODUCTS: AT LEAST FIVE DIFFERENT ISOFORMS; GGFBP1,			
CC	GGFBP2, GGFBP3, GGFBP4 AND GGFBP5 (SHOWN HERE); ARE PRODUCED			
CC	BY ALTERNATIVE SPLICING.			
CC	- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.			
CC	- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.			
CC	- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.			
DR	EMBL: U12259; AAA30540.1; -.			
DR	HSPF: Q12784; IHRE.			
DR	InterPro: IPR000561; EGF-like.			
DR	InterPro: IPR003598; Ig_C2.			
DR	InterPro: IPR003006; Ig_MHC.			
DR	Pfam: PF000008; EGF_1.			
DR	Pfam: PF00047; Ig_1.			
DR	SMART: SM00181; EGF_1.			
DR	SMART: SM00408; IG_C2_1.			
DR	PROSITE: PS00022; EGF_1; UNKNOWN_1.			
DR	PROSITE: PS01186; EGF_2; FALSE_NEG.			
DR	Growth factor: EGF-like domain; Immunoglobulin domain;			
DR	Alternative splicing.			
FT	PROPEP	1	19	
FT	CHAIN	20	241	
FT				BY SIMILARITY.
FT				NEUREGULIN-1, GLIAL GROWTH FACTOR 5
FT				ISOFORM.
FT	DOMAIN	50	119	IG-LIKE C2-TYPE DOMAIN.
FT	DOMAIN	165	177	SER/TIR-RICH.
FT	DOMAIN	178	222	EGF-LIKE DOMAIN.
FT	DISULEFD	182	222	BY SIMILARITY.
FT	DISULEFD	182	222	BY SIMILARITY.
FT	DISULEFD	190	210	BY SIMILARITY.
FT	DISULEFD	212	221	BY SIMILARITY.
FT	VARSPLIC	134	226	EITGMASMETAVYSESPIRI -> GKRCCLRAISQSLR
FT				GVIKVGCHT (IN ISOFORM GGFBP1).
FT	VARSPLIC	157	241	MISSING (IN ISOFORM GGFBP1).
FT	SEQUENCE	241 AA;	25955 MW;	BS71297EBDA9J76 CRC64;

OY	1	HFKRCRDRLAYCINLNGECVETITLGSHKH-CRCKEGYGVRCDQFL	47
		: ::  ::   :: :	:: :: :: :: :: :
Dd	178	HLVACCAKEKTEFCVNGSECCMWADLSNPSRYLCKCPNETGTGRQNYY	225
RESULT	5		
Q96IB3		PRELIMINARY;	PRT; 296 AA.
AC	Q96IB3;		
DT	01-DEC-2001 (TREMBLrel_19,	Created)	
DT	01-DEC-2001 (TREMBLrel_19,	Last sequence update)	
DT	01-DEC-2001 (TREMBLrel_19,	Last annotation update)	
DE	NEUROREGULIN 1.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	TISSUE=BRAIN, AND NEUROBLASTOMA;		
RA	Strassberg R.;		
RL	Submitted (MAY-2001) to the EMBL/Genbank/DDBJ databases.		
DR	EMBL: BC007675; AAH07675.1; -		
SQ	SEQUENCE 296 AA; 31659 MW; 0F5C48CB65B6649 CRC64;		

[illegible]

RESULT	6			
ID	09ESAB	PRELIMINARY;	PRT;	111 AA.
AC	09ESAB:			
DT	01-MAR-2001 (TREMblrel. 16, Created)			
DT	01-MAR-2001 (TREMblrel. 16, Last sequence update)			
DT	01-DEC-2001 (TREMblrel. 19, Last annotation update)			
DE	SMDF NEUREGULIN BETA 2 (FRAGMENT).			
GN	NRG1.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus			
OX	NCBI_Taxid=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SPRAGUE-DAWLEY;			
RC	TISSUE=AXOTOMIZED LUMBAR DORSAL ROOT GANGLION/SPINAL CORD;			
RC	CARROLL S.L., ANDERSON K.D., FROHNEIT P.W.;			
RA	"Structural and Functional Diversity of SMDF Neuregulin Splice			
RT	Variants Expressed in the Adult Rat Nervous System."			
RT	Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.			
RL	EMBL: AF194441; AAC28430.1; -			
DR	HSSP: Q12784; IRR.			
DR	InterPro: IPR000561; EGF-like.			
DR	InterPro: IPR000866; ER-target.			
DR	Pfam: PF00008; EGF. 1.			
DR	SMART: SM00181; EGF. 1.			
DR	SMART: SM00014; EGF-like. 1.			
DR	PROSITE: PS00014; ER_TARGET; UNKNOWN_1.			
FT	NON_TER	1		
FT	NON_TER	1		
FT	NON_TER	111	111	
SEQUENCE	111 AA, 1219 MW, CGBBB70584C9F8C CRC64;			

Query Match	37.7%	Score 104.5;	DB 6;	Length 241;	
Best Local Similarity	31.2%;	Pred. No. 9.7e-07;			
Matches	15;	Conservative	14;	Mismatches	18;
				Indels	1;
				Gaps	1;

---

Query Match	37.0%	Score 102.5;	DB 11;	Length 111;	
Best Local Similarity	31.2%;	Pred. No. 8.1e-07;			
Matches	15;	Conservative	13;	Mismatches	19;
				Indels	1;
				Gaps	1;





GN NRGL.  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SPRAGUE-DAWLEY;  
RA Carroll S.L., Stonecypher M.S., Anderson K.D., Pearson R.J. Jr.,  
RT Frohner P.W.,  
\*Structural and Functional Diversity of Glial Growth Factor Isoforms  
Expressed in Regenerating Peripheral Nerve and Associated Neurons.\*;  
RL Submitted (Oct-1999) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AF194996; AAC28450.1; -  
DR HSSP; Q12784; IHRF.  
DR InterPro; IPR000561; EGF\_1like.  
DR InterPro; IPR003599; IG.  
DR InterPro; IPR003598; IG\_C2.  
DR InterPro; IPR003600; IG\_1like.  
DR InterPro; IPR003006; IG\_MHC.  
DR Pfam; PF00047; IG; 1.  
DR SMART; SM00181; EGF; 1.  
DR SMART; SM00409; EGF\_1like; 1.  
DR SMART; SM00408; IGC2; 1.  
DR SMART; SM00410; IG\_1like; 1.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
KW Immunoglobulin domain.  
FT NON\_TER  
SQ SEQUENCE 323 AA; 35358 MW; C7DF153A939A80C8 CRC64;

Query Match 37.0%; Score 102.5; DB 11; Length 323;  
Best Local Similarity 31.2%; Pred. No. 2.5e-06;  
Matches 15; Conservative 13; Mismatches 19; Indels 1; Gaps 1;

OY 1 HFKPRDKLALYCLNDGECFYETLTGSHK-CRCKEGYGVRCDOFL 47  
DB 260 HLKCAKEKTEFCVNGECFTVKDLSNPSRYLCKCPNEFTGDRCONYV 307

RESULT 11  
OQESAI PRELIMINARY; PRT; 342 AA.  
AC OQESAI;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update);  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE GLIAL GROWTH FACTOR GGF BETA 4 (FRAGMENT).  
GN NRGL.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SPRAGUE-DAWLEY;  
RA Carroll S.L., Stonecypher M.S., Anderson K.D., Pearson R.J. Jr.,  
RT Frohner P.W.,  
\*Structural and Functional Diversity of Glial Growth Factor Isoforms  
Expressed in Regenerating Peripheral Nerve and Associated Neurons.\*;  
RL Submitted (Oct-1999) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AF194997; AAC28451.1; -  
DR HSSP; Q12784; IHRF.  
DR InterPro; IPR000561; EGF\_1like.  
DR InterPro; IPR003599; IG.  
DR InterPro; IPR003598; IG\_C2.  
DR InterPro; IPR003600; IG\_1like.  
DR Pfam; PF00047; IG; 1.  
DR SMART; SM00181; EGF; 1.  
DR SMART; SM00001; EGF\_1like; 1.  
DR SMART; SM00409; IG; 1.

DR SMART; SM00408; IGC2; 1.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
KW Immunoglobulin domain.  
FT NON\_TER  
SQ SEQUENCE 342 AA; 37836 MW; 8BE36FC836553124 CRC64;

Query Match 37.0%; Score 102.5; DB 11; Length 342;  
Best Local Similarity 31.2%; Pred. No. 2.7e-06;  
Matches 15; Conservative 13; Mismatches 19; Indels 1; Gaps 1;

OY 1 HFKPRDKLALYCLNDGECFYETLTGSHK-CRCKEGYGVRCDOFL 47  
DB 260 HLKCAKEKTEFCVNGECFTVKDLSNPSRYLCKCPNEFTGDRCONYV 307

RESULT 12  
OQESBI PRELIMINARY; PRT; 700 AA.  
AC OQESBI;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE SMOF NEUREGULIN BETA 1A.  
GN NRGL.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SPRAGUE-DAWLEY;  
RA Carroll S.L., Anderson K.D., Frohner P.W.,  
RT Frohner P.W.,  
\*Structural and Functional Diversity of SMOF Neuregulin Splice  
Variants Expressed in the Adult Rat Nervous System.\*;  
RL Submitted (Oct-1999) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AF194438; AAC28427.1; -  
DR HSSP; Q12784; IHRF.  
DR InterPro; IPR000561; EGF\_1like.  
DR InterPro; IPR002154; Neuregulin.  
DR InterPro; IPR002114; PRTS\_HPR\_ser.  
DR Pfam; PF00008; EGF; 1.  
DR Pfam; PF02158; Neuregulin; 1.  
DR PRINTS; PRO1089; NEUREGULIN.  
DR SMART; SM00181; EGF; 1.  
DR SMART; SM00001; EGF\_1like; 1.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
DR PROSITE; PS00589; PRTS\_HPR\_ser; UNKNOWN\_1.  
SQ SEQUENCE 700 AA; 76386 MW; 2F811B1B7ECC49DA CRC64;

Query Match 37.0%; Score 102.5; DB 11; Length 700;  
Best Local Similarity 31.2%; Pred. No. 5.7e-06;  
Matches 15; Conservative 13; Mismatches 19; Indels 1; Gaps 1;

OY 1 HFKPRDKLALYCLNDGECFYETLTGSHK-CRCKEGYGVRCDOFL 47  
DB 234 HLKCAKEKTEFCVNGECFTVKDLSNPSRYLCKCPNEFTGDRCONYV 281

RESULT 13  
OQESAS PRELIMINARY; PRT; 782 AA.  
AC OQESAS;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE GLIAL GROWTH FACTOR BETA 1A (FRAGMENT).  
GN NRGL.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=SPINAL CORD/BRAIN STEM;  
 RA Carroll S.L., Stonecypher M.S., Anderson K.D., Pearson R.J. Jr.,  
 RT Frohner P.W.;  
 RT "Structural and Functional Diversity of Glial Growth Factor Isoforms  
 RT Expressed in Regenerating Peripheral Nerve and Associated Neurons";  
 RT Submitted (Oct-1999) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF194993; AAG28433.1; -  
 DR HSSP: Q12784; 1HRE.  
 DR InterPro: IPR000561; EGF\_1like.  
 DR InterPro: IPR003598; 1g.  
 DR InterPro: IPR003598; 1g.  
 DR InterPro: IPR003006; 1g.  
 DR InterPro: IPR002154; Neuregulin.  
 DR InterPro: IPR002114; Pts\_Hpr\_ser.  
 DR Pfam: PF00047; 1g.  
 DR Pfam: PF02158; Neuregulin; 1.  
 DR PRINTS: PRO1089; NEUREGULIN.  
 DR SMART: SM00181; EGF; 1.  
 DR SMART: SM00001; EGF\_1like; 1.  
 DR SMART: SM00409; IG; 1.  
 DR SMART: SM00408; IGC2; 1.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_1.  
 DR PROSITE: PS00589; Pts\_Hpr\_ser; UNKNOWN\_1.  
 DR KEGG: K01001; Immunoglobulin domain.  
 DR NON\_TER 1  
 SQ SEQUENCE 782 AA; 86036 MW; F6174A68F4E27BDE CRC64;

Query Match 37.0%; Score 102.5; DB 11; Length 782;  
 Best Local Similarity 31.2%; Pred. No. 6.5e-06;  
 Matches 15; Conservative 13; Mismatches 19; Indels 1; Gaps 1;

OY 1 HKPKRDKLAVCLNDGCEVFIEITLGSKH-CRCCKEGYGVRCDFL 47  
 Db 316 HLKCAKEKTECVANGGCEFTVKDLSNPSRYLCKPCNFTGDRCONYV 363

RESULT 14  
 O91M20 PRELIMINARY; PRT; 89 AA.  
 AC O91M20;  
 DT 01-DEC-2001 (TREMblrel. 19, Created)  
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE LSDV016 EGF-LIKE GROWTH FACTOR.  
 GN LSDV016.  
 OS Lumpy skin disease virus.  
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Capripoxvirus.  
 OX NCBI\_TaxID=59509;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NEERHLING 2490;  
 RX MEDLINE=2139495; PubMed=11435593;  
 RA Tulman E.R., Afonso C.L., Lu Z., Zsak L., Kutish G.F., Rock D.L.;  
 RT "Genome of lumpy skin disease virus";  
 RT J. Virol. 75:7122-7130(2001).  
 RL [2]  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NEERHLING 2490;  
 RA Tulman E.R., Afonso C.L., Lu Z., Zsak L., Kutish G.F., Rock D.L.;  
 RT Submitted (Aug-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF325528; AAK84977.1; -  
 SQ SEQUENCE 89 AA; 10646 MW; ID5F3FD7D06174E0 CRC64;

Query Match 35.7%; Score 99; DB 12; Length 89;  
 Best Local Similarity 43.9%; Pred. No. 2e-06;  
 Matches 18; Conservative 6; Mismatches 13; Indels 4; Gaps 1;

OY 7 DKDLAYCLNDGCEVFIEITLGSKH---HRCCKEGYGVRC 43  
 Db 40 DKSLNFCNLGCTCYKTFITLSSYNKKPLMFCRCCKLGYGVRC 80

RESULT 15  
 ID 001768 PRELIMINARY; PRT; 2192 AA.  
 AC 001768;  
 DT 01-JUL-1997 (TREMblrel. 04, Created)  
 DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE HYPOTHETICAL 242.7 KDA PROTEIN.  
 GN T1E3.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA None;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium";  
 RT Science 282:2012-2018(1998).  
 RL [2]  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA Du Z., Le T.T.;  
 RT "The sequence of C. elegans cosmid T1E3";  
 RT submitted (May-1997) to the EMBL/Genbank/DBJ databases.  
 RL [3]  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA Waterston R.;  
 RT "Direct Submission";  
 RT submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.  
 DL EMBL: AF003133; AAB54138.2; -  
 DR HSSP: Q07954; 1CR8.  
 DR InterPro: IPR000561; EGF\_1like.  
 DR InterPro: IPR002049; Laminin\_EGF.  
 DR InterPro: IPR000033; LDL\_receptor.  
 DR InterPro: IPR002172; LDL\_receptor\_A.  
 DR Pfam: PF00008; EGF; 8.  
 DR Pfam: PF00057; LDL\_receptor\_a; 18.  
 DR Pfam: PF00058; LDL\_receptor\_b; 8.  
 DR PRINTS: PRO0011; EGF\_LAMININ.  
 DR PRINTS: PRO0261; LDLRECEPTOR.  
 DR SMART: SM00181; EGF; 23.  
 DR SMART: SM00192; LDLa; 18.  
 DR SMART: SM00135; LY; 7.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_9.  
 DR PROSITE: PS01186; EGF\_2; 7.  
 DR PROSITE: PS01209; LDLRA\_1; 10.  
 DR PROSITE: PS00068; LDLRA\_2; 12.  
 KW EGF-like domain; Glycoprotein; Hypothetical protein.  
 SQ SEQUENCE 2192 AA; 242666 MW; F4762A5EBCA45BDA CRC64;

Query Match 33.4%; Score 92.5; DB 5; Length 2192;  
 Best Local Similarity 42.9%; Pred. No. 0.00049;  
 Matches 18; Conservative 8; Mismatches 7; Indels 9; Gaps 3;  
 OY 5 CRDKLAVCLNDGCEVFIEITLGSKHRCCKEGYGVRCDFL 45  
 Db 1918 CDD---YCTNNSKC---TTINGTHFEDCKKGFGRKCEQ 1951

Search completed: September 13, 2002, 08:55:27  
 Job time: 354 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 08:49:53 ; Search time 10.36 Seconds

(without alignments)  
175.658 Million cell updates/sec

Title: US-09-877-665-4

Perfect score: 277  
1 HFRPCRKDLAYCLNDGECE.....SHKRCCKEGYGVRCDFL 47

Sequence: 1 HFRPCRKDLAYCLNDGECE.....SHKRCCKEGYGVRCDFL 47

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Database: SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	277	100.0	713	1	NRG3_MOUSE
2	277	100.0	720	1	NRG3_HUMAN
3	126.5	45.7	677	1	NRG1_XENLA
4	113.5	41.0	639	1	NRG1_HUMAN
5	110	39.7	115	1	NRG4_MOUSE
6	104.5	37.7	296	1	SMDE_HUMAN
7	104	37.5	756	1	NRG2_MOUSE
8	102.5	37.0	662	1	NRG1_RAT
9	98.5	35.6	602	1	NRG1_CHICK
10	90	32.5	850	1	NRG2_HUMAN
11	90	32.5	868	1	NRG2_RAT
12	89.5	32.3	169	1	EREG_MOUSE
13	85	30.7	80	1	GRFA_SEVKA
14	83	30.0	2531	1	NRG1_MOUSE
15	82.5	29.8	177	1	NRG1_MOUSE
16	82.5	29.8	178	1	NRG1_MOUSE
17	82	29.6	85	1	GRFA_MYXL
18	82	29.6	230	1	SPIT_DROME
19	80.5	29.1	125	1	NRG1_MOUSE
20	80.5	29.1	178	1	NRG1_MOUSE
21	79.5	28.7	140	1	NRG1_MOUSE
22	79.5	28.7	142	1	NRG1_MOUSE
23	79.5	28.7	159	1	NRG1_MOUSE
24	79.5	28.7	159	1	NRG1_MOUSE
25	79	28.5	2531	1	NRG1_MOUSE
26	78	28.2	1207	1	NRG1_MOUSE
27	77.5	28.0	714	1	NRG1_MOUSE
28	77.5	28.0	722	1	NRG1_MOUSE
29	76	27.4	484	1	NRG1_MOUSE
30	76	27.4	1217	1	NRG1_MOUSE
31	76	27.4	2139	1	NRG1_MOUSE
32	75	27.1	611	1	NRG1_MOUSE
33	74.5	26.9	603	1	NRG1_MOUSE

34	74.5	26.9	723	1	DL11_HUMAN	O00548 homo sapien
35	74.5	26.9	1429	1	L112_CAEL	P14585 caenorhabd
36	74	26.7	551	1	LEM2_RABIT	P27113 oryctolagus
37	74	26.7	610	1	LEM2_HUMAN	P16581 homo sapien
38	74	26.7	2871	1	FBN1_BOVIN	P98133 bos taurus
39	74	26.7	2871	1	FBN1_HUMAN	P35555 homo sapien
40	74	26.7	2871	1	FBN1_MOUSE	Q61554 mus musculu
41	74	26.7	2871	1	FBN1_MOUSE	Q9TV36 sus scrofa
42	73	26.4	294	1	GRK_DROME	P42287 drosophila
43	73	26.4	409	1	MGK_DROME	P79385 sus scrofa
44	73	26.4	1964	1	NTC4_MOUSE	P31695 mus musculu
45	73	26.4	2524	1	NOTC_XENLA	P21783 xenopus lae

## ALIGNMENTS

RESULT 1  
NRG3\_MOUSE STANDARD: PRT; 713 AA.  
AC 035181;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Pro-neuregulin-3 precursor (Pro-NRG3) [Contains: Neuregulin-3 (NRG-3)]  
DE NRG3.  
GN NRG3.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
ON NCBI\_Taxid=10090;  
RX [1]  
RP TISSUE: FROM N.A.  
RC TISSUE: Brain;  
RX MEDLINE-97420720; PubMed-9275162;  
RA Zhang D., Sliwkowski M.X., Mark M., Frantz G., Akita R., Sun Y., Hillan K., Crowley C., Brush J., Godowski P.J.;  
RT 'Neuregulin-3 (NRG3): a novel neural tissue-enriched protein that binds and activates ErbB4'.  
RL Proc. Natl. Acad. Sci. U.S.A. 94:9562-9567(1997).  
CC -1- FUNCTION: DIRECT LIGAND FOR THE ERBB4 TYROSINE KINASE RECEPTOR. BINDING RESULTS IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND ACTIVATION OF THE RECEPTOR. DOES NOT BIND TO THE EGF RECEPTOR, ERBB2 OR ERBB3 RECEPTORS.  
CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS A PROPELTYCALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN SYMPATHETIC, MOTOR, AND SENSORY NEURONS.  
CC -1- DEVELOPMENTAL STAGE: DETECTED AS EARLY AS E11. IN E13 EMBRYOS, DETECTED MAINLY IN THE NERVOUS SYSTEM. IN E16 EMBRYOS, DETECTED IN THE BRAIN, SPINAL CORD, TRIGEMINAL, VESTIBULAR-COCHLEAR, AND SPINAL GANGLIA. IN ADULTS, EXPRESSED IN SPINAL CORD, AND NUMEROUS BRAIN REGIONS.  
CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN DIMERIZATION (BY SIMILARITY).  
CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE DOMAIN (BY SIMILARITY).  
CC -1- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR FORM (BY SIMILARITY).  
CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.  
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.  
-----  
This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: AF010130; AAB70914.1; ..  
 CC MGD: MGI:1097165; Nf93.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR002154; Neuregulin.  
 DR Pfam: PF00008; EGF; 1.  
 DR Pfam: PF02158; Neuregulin; 1.  
 DR SMART: SM00181; EGF; 1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 KM Growth factor; EGF-like domain; Transmembrane; Multigene family.  
 FT CHAIN 1 713 PRO-NEUREGULIN-3, MEMBRANE-BOUND FORM.  
 FT CHAIN 1 362 NEUREGULIN-3.  
 FT DOMAIN 1 362 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 363 383 INTERNAL SIGNAL SEQUENCE (POTENTIAL).  
 FT DOMAIN 384 713 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 105 287 SER/THR-RICH.  
 FT DOMAIN 288 331 EGF-LIKE.  
 FT DOMAIN 13 21 POLY-ALA.  
 FT DOMAIN 26 34 POLY-ALA.  
 FT DOMAIN 127 135 POLY-THR.  
 FT DOMAIN 250 253 POLY-ALA.  
 FT DOMAIN 254 263 POLY-SER.  
 FT DOMAIN 264 267 POLY-THR.  
 FT DISULFID 292 306 BY SIMILARITY.  
 FT DISULFID 300 319 BY SIMILARITY.  
 FT DISULFID 321 330 BY SIMILARITY.  
 SQ SEQUENCE 713 AA; 77369 MW; 9F7D1D5E7FC8DC60 CRC64;

Query Match 100.0%; Score 277; DB 1; Length 713;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-26;  
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFKPCRDKLAVCLNDGECFVETLTGSHKRCRKEGYGVRCDOFL 47  
 DB 288 HFKPCRDKLAVCLNDGECFVETLTGSHKRCRKEGYGVRCDOFL 334

RESULT 2  
 NRG3\_HUMAN STANDARD; PRT; 720 AA.  
 AC P56975;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Pro-neuregulin-3 precursor (Pro-NRG3) [Contains: Neuregulin-3 (NRG-3)].  
 GN NRG3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fetal brain;  
 RX MEDLINE=97420720; PubMed=9275162;  
 RA Zhang D., Sliwkowski M.X., Mark M., Frantz G., Akita R., Sun Y.,  
 RA Hillan K., Crowley C., Brush J., Godowski P.J.;  
 RT "Neuregulin-3 (NRG3): a novel neural tissue-enriched protein that  
 binds and activates ErbB4.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:9562-9567(1997).  
 CC -1- FUNCTION: DIRECT LIGAND FOR THE ERBB4 TYROSINE KINASE RECEPTOR.  
 CC BINDING RESULTS IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND  
 CC ACTIVATION OF THE RECEPTOR. DOES NOT BIND TO THE EGF RECEPTOR.  
 CC ERBB2 OR ERBB3 RECEPTORS.  
 CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS  
 CC A PROTEOLITICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE  
 CC MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN MOST REGIONS OF THE BRAIN  
 CC WITH THE EXCEPTION OF CORPUS CALLOSUM. EXPRESSED AT LOWER LEVEL IN

TESTIS. NOT DETECTED IN HEART, PLACENTA, LUNG, LIVER, SKELETAL  
 MUSCLE, KIDNEY, PANCREAS, SPLEEN, THYMUS, PROSTATE, OVARY, SMALL  
 INTESTINE, COLON AND PERIPHERAL BLOOD LEUKOCYTES.  
 CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION  
 CC OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE  
 CC PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN  
 CC DIMERIZATION (BY SIMILARITY).  
 CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE  
 CC DOMAIN (BY SIMILARITY).  
 CC -1- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE  
 CC EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR  
 CC FORM (BY SIMILARITY).  
 CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.  
 DR MIM: 605533; ..  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR002154; Neuregulin.  
 DR Pfam: PF02158; Neuregulin; 1.  
 DR SMART: SM00181; EGF; 1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 KM Growth factor; EGF-like domain; Transmembrane; Multigene family.  
 FT CHAIN 1 720 PRO-NEUREGULIN-3, MEMBRANE-BOUND FORM.  
 FT CHAIN 1 359 NEUREGULIN-3.  
 FT DOMAIN 1 360 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 361 381 INTERNAL SIGNAL SEQUENCE (POTENTIAL).  
 FT DOMAIN 382 720 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 105 285 SER/THR-RICH.  
 FT DOMAIN 286 329 EGF-LIKE.  
 FT DOMAIN 5 8 POLY-ALA.  
 FT DOMAIN 13 21 POLY-ALA.  
 FT DOMAIN 26 34 POLY-ALA.  
 FT DOMAIN 127 135 POLY-THR.  
 FT DOMAIN 252 260 POLY-SER.  
 FT DOMAIN 265 265 POLY-THR.  
 FT DISULFID 290 304 BY SIMILARITY.  
 FT DISULFID 298 317 BY SIMILARITY.  
 FT DISULFID 319 328 BY SIMILARITY.  
 SQ SEQUENCE 720 AA; 77900 MW; AAD6F10DB95A693 CRC64;

Query Match 100.0%; Score 277; DB 1; Length 720;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-26;  
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFKPCRDKLAVCLNDGECFVETLTGSHKRCRKEGYGVRCDOFL 47  
 DB 286 HFKPCRDKLAVCLNDGECFVETLTGSHKRCRKEGYGVRCDOFL 332

RESULT 3  
 NRG1\_XENLA STANDARD; PRT; 677 AA.  
 AC O93383; O9W6N0;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Pro-neuregulin-1 precursor (Pro-NRG1) [Contains: Neuregulin-1].  
 GN NRG1.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;  
 OC Xenopodidae; Xenopus.  
 OC NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM ALPHA1), AND ALTERNATIVE SPLICING.  
 RX MEDLINE=98352126; PubMed=9685585;  
 RA Yang J.F., Zhou H., Pun S., Ip N.Y., Peng H.B., Tsai K.W.K.;  
 RT "Cloning of cDNAs encoding xenopus neuregulin: expression in myotomal  
 muscle during embryo development.";  
 RL Brain Res. Mol. Brain Res. 58:59-73(1998).

FN	[2]	SEQUENCE FROM N. A. (ISOFORM CRD).
RP		MEDLINE=99316087; PubMed=10383827;
RX		Yang J.F., Zhou H., Choi R.C., Ip N.Y., Peng H.B., Tsim K.W.K.;
RA		"A cysteine-rich form of Xenopus neuroguin induces the expression of
RT		acetylcholine receptors in cultured myotubes.";
RL		Mol. Cell. Neurosci. 13:415-429(1999).
CC	-1-	FUNCTION: DIRECT LIGAND FOR THE ERBB TYROSINE KINASE RECEPTORS.
CC	-1-	INDUCES EXPRESSION OF ACETYLCHOLINE RECEPTOR IN SYNAPTIC NUCLEI.
CC	-1-	SUBCELLULAR LOCATION: EXISTS AS A TYPE I MEMBRANE PROTEIN AND AS A
CC	-1-	PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE MEMBRANE-
CC	-1-	BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).
CC	-1-	ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; ALPHA1 (SHOWN HERE) AND
CC	-1-	CRD/CRD-NR01 (CYSTEINE RICH DOMAIN); ARE PRODUCED BY ALTERNATIVE
CC	-1-	SPLICING. ISOFORMS HAVE ALPHA OR BETA-TYPE EGF-LIKE DOMAINS.
CC	-1-	TISSUE SPECIFICITY: ISOFORM ALPHA1 IS EXPRESSED IN BRAIN AND
CC	-1-	MUSCLE. ISOFORM CRD IS EXPRESSED IN BRAIN AND SPINAL CORD, BUT
CC	-1-	AT VERY LOW LEVEL IN MUSCLE.
CC	-1-	DEVELOPMENTAL STAGE: STRONG EXPRESSION IN DEVELOPING BRAIN AND
CC	-1-	SPINAL CORD OF THE EMBRYO. ALSO EXPRESSED IN THE MYOTOMAL MUSCLE.
CC	-1-	DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION
CC	-1-	OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE
CC	-1-	PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN
CC	-1-	DIMERIZATION (BY SIMILARITY).
CC	-1-	DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE
CC	-1-	DOMAIN.
CC	-1-	PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE
CC	-1-	EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR
CC	-1-	FORM.
CC	-1-	PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY
CC	-1-	SIMILARITY).
CC	-1-	SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC	-1-	SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC	-1-	SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
CC	-1-	THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
CC	-1-	BETWEEN THE SWISS INSTITUTE OF BIOFORMATICS AND THE EMBL OUTSTATION -
CC	-1-	THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
CC	-1-	USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
CC	-1-	MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
CC	-1-	ENTITIES REQUIRES A LICENSE AGREEMENT (SEE <a href="http://www.isb-sdb.ch/announce/">http://www.isb-sdb.ch/announce/</a>
CC	-1-	OR SEND AN EMAIL TO <a href="mailto:license@sdb-sdb.ch">license@sdb-sdb.ch</a> ).
DR		EMBL; AF076618; AAC26804.1; -
DR		EMBL; AF142632; AAD33893.1; -
DR		HSSP; Q12784; IHRE.
DR		InterPro: IPR0000561; EGF-like.
DR		InterPro: IPR003006; Ig_MHC.
DR		InterPro: IPR003598; Ig_C2.
DR		InterPro: IPR002154; Neuregulin.
DR		pfam; PF000008; EGF; 1.
DR		pfam; PF000047; Ig; 1.
DR		pfam; PF02158; Neuregulin; 1.
DR		PRINTS: PR01089; NEUREGULIN.
DR		SMART; SM00181; EGF; 1.
DR		SMART; SM00408; IGC2; 1.
DR		PROSITE; PS00022; EGF_1; 1.
DR		PROSITE; PS01186; EGF_2; 1.
KW		Growth factor; EGF-like domain; Immunoglobulin domain; Glycoprotein;
KW		Transmembrane; Alternative splicing.
FT	CHAIN	1 259
FT	CHAIN	1 677
FT	DOMAIN	1 260
FT	TRANSMEM	261 280
FT	DOMAIN	281 677
FT	DOMAIN	50 119
FT	DOMAIN	188 232
FT	DISULFID	57 116
FT	DISULFID	192 206
FT	DISULFID	200 220
FT	DISULFID	222 231
FT	DOMAIN	1 25
FT		LYS-RICH.

FT FT CARBOHYD 124 124 (POTENTIAL).  
 FT CARBOHYD 130 130 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT VARSPLIC 1 136 MAEKKEVREKSKKGGKDRGKAKAGSGDGAAPRIKE  
 FT ITO5OEBKRLKLOAVSEPSLKFNFNGKEKIGANK  
 FT NOTVAVNTVIVR -> MSEDYAEGLONGCESSDPPSAB  
 FT LNEESMPETOEEHTHGTGATGACCCLEADRRLCLN  
 FT SKRCIIPILACISLCIAGIKAFVYKIFEYOSPHLD  
 FT PCHRGODLLTYDTPASTVPSVYRLPIYITTTSKAAYT  
 FT EFRGSLPTE (IN ISOFORM CRD).  
 FT KCGFVGARCTEYDPLRVNSEKHLGIEFNE -> PNEFTGD  
 FT RCONVYMASFYK (IN ISOFORM CRD).  
 SQ SEQUENCE 677 AA; 75794 MW; 49279EBF5BAE396F CRC64;  
 Query Match 45.7%; Score 126.5; DB 1; Length 677;  
 Best Local Similarity 43.5%; Pred. No. 6, 6e-08;  
 Matches 20; Conservative 10; Mismatches 15; Indels 1; Gaps 1.  
 QY 1 HEPKCRKDLAYCLNDECFVETLTGSKHK-CRCGEYGVAVCDQ 45  
 Db 188 HLIKSDKREKTYCVANGCEYVINGLTSSNQFMCKCKPGFTGARCTE 233  
 RESULT 4  
 NRGI\_HUMAN STANDARD; PRT; 639 AA.  
 ID NRGI\_HUMAN  
 AC 002299; 002299; 002299; 012779; 012780; 012781; 012782; 012783;  
 AC 012784; 007110; 007111; 009693; 014667;  
 DT 16-OCT-2001 (Rel. 40; Created)  
 DT 16-OCT-2001 (Rel. 40; Last sequence update)  
 DT 16-OCT-2001 (Rel. 40; Last annotation update)  
 DE Pro-neuregulin-1 precursor (Pro-NG1) [Contents: Neuregulin-1 (Neu  
 DE differentiation factor) (Heregulin) (HRG) (Breast cancer cell  
 DE differentiation factor p45) (Acetylcholine receptor inducing activity)  
 DE (AR4) (Sensory and motor neuron-derived factor) (Glial growth  
 DE factor)].  
 GN NRGI OR HGL OR NDF OR HRGA OR GGF OR SMDF.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_Taxid:9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS ALPHA; BETA1; BETA2 AND BETA3), AND  
 RP PARTIAL SEQUENCE.  
 RX MEDLINE=92271253; PubMed=1350381;  
 RA Holmes W.E., Sliakowski M.X., Akita R.W., Henzel W.J., Lee J.,  
 RA Park J.W., Yansura D., Abadi N., Raab H., Lewis G.D., Shepard H.M.,  
 RA Kuang W.-J., Wood W.I., Goeddel D.V., Vanden R.L.,  
 RT "Identification of heregulin, a specific activator of p185erbB2";  
 RL Science 256:1205-1210(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS ALPHA1A; ALPHA2B; ALPHA3; BETA1; BETA2  
 RP AND BETA3).  
 RP TISSUE-pituitary, and kidney adenocarcinoma;  
 RX MEDLINE=9415863; PubMed=7509448;  
 RA Wen D., Suggs S.V., Katunagaran D., Liu N., Cupples R.L., Luo Y.,  
 RA Janssen S.-Y., Ben-Baruch N., Trollinger D.B., Jacobsen V.L.,  
 RA Meng S.-Y., Lu H.-S., Lu S., Chang D., Yang W., Yanigahara D.,  
 RA Koski R.A., Yarden Y.,  
 RT "Structural and functional aspects of the multiplicity of Neu  
 RT differentiation factors";  
 RL Mol. Cell. Biol. 14:1909-1919(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM ALPHA).  
 RX MEDLINE=92208945; PubMed=1348215;  
 RA Pelves E., Bacus S.S., Koski R.A., Lu H.S., Wen D., Ogden S.G.,  
 RA Levy R.B., Yarden Y.,  
 RT "Isolation of the neu/HER-2 stimulatory ligand: a 44 kd glycoprotein  
 RT that induces differentiation of mammary tumor cells";  
 RL Cell 69:205-216(1992).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORMS BETA3 AND GGF2).





DR Pfam: PF02158; Neuregulin-1.  
 DR PRINTS: PRO1089; NEUREGULIN.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 DR Growth factor; EGF-like domain; Immunoglobulin domain; Glycoprotein; 3D-structure;  
 DR Transmembrane; Multigene family; Polymorphism; 3D-structure;  
 DR Alternative splicing; Chromosomal translocation.  
 KW  
 Query Match 41.0%; Score 113.5; DB 1; Length 639;  
 Best Local Similarity 34.8%; Pred. No. 2,3e-06;  
 Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;  
 Oy 1 HFKRCRDKLAVCLNDGCEYIEITLGSKH-CCKCKEYQGVRCQ 45  
 Db 177 HLVCAERKTECVNGGCEVFKDLSPSRYLCKQPGFTGACRTE 222  
 RESULT 5  
 NRGA\_MOUSE STANDARD; PRT; 115 AA.  
 ID 09MTX4;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Pro-neuregulin-4, short isoform (Pro-NRG4) [contains: Neuregulin-4  
 (NRG-4)].  
 GN NRG4.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_Taxid=10090;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Liver;  
 RX MEDLINE=99276098; PubMed=10348342;  
 RA Harari D., Tzahar E., Romano J., Shelly M., Pierce J.H., Andrews G.C.,  
 RT Yarden Y.;  
 RA "Neuregulin-4: a novel growth factor that acts through the ErbB-4  
 Receptor tyrosine kinase.";  
 RL Oncogene 18:2681-2689(1999).  
 CC -1- FUNCTION: LOW AFFINITY LIGAND FOR THE ERBB4 TYROSINE KINASE  
 RECEPTOR. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,  
 RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND  
 ACTIVATION OF THE ERBB RECEPTORS. DOES NOT BIND TO THE ERBB1,  
 ERBB2 AND ERBB3 RECEPTORS.  
 CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS  
 A PROTOOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE  
 MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).  
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS MAY BE PRODUCED BY  
 ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PANCREAS; WEAKLY EXPRESSED  
 IN MUSCLE.  
 CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION  
 OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE  
 PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN  
 DIMERIZATION (BY SIMILARITY).  
 CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE  
 DOMAIN (BY SIMILARITY).  
 CC -1- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE  
 EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR  
 FORM (BY SIMILARITY).  
 CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY  
 SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.  
 CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL: AF083067; AAD21874.1; -  
 DR MGD; MG1:1933833; Nrg4.  
 DR InterPro: IPR000561; EGF\_1like.  
 DR InterPro: IPR001336; EGF\_1.  
 DR Pfam: PF00008; EGF; 1.  
 DR PRINTS: PRO0009; EGF; 1.  
 DR SMART: SM00181; EGF; 1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; FALSE\_NEG.  
 DR Growth factor; EGF-like domain; Glycoprotein; Transmembrane;  
 KW Multigene family; Alternative splicing.  
 KW CHAIN 1 115  
 FT CHAIN 1 61  
 FT DOMAIN 1 62  
 FT TRANSMEM 63 83  
 FT DOMAIN 84 115  
 FT DOMAIN 5 46  
 FT DISULFID 9 23  
 FT DISULFID 17 34  
 FT DISULFID 36 45  
 FT CARBOHYD 39 39  
 FT CARBOHYD 60 60  
 SQ SEQUENCE 115 AA; 12743 MW; 989A1E376F857B49 CRC64;  
 Query Match 39.7%; Score 110; DB 1; Length 115;  
 Best Local Similarity 42.2%; Pred. No. 1,2e-06;  
 Matches 19; Conservative 8; Mismatches 16; Indels 2; Gaps 1;  
 Oy 1 HFKRCRDKLAVCLNDGCEYIEITLGSKHCKCKEYQGVRCQ 45  
 Db 5 HEQCPGRHRSFCLNGICGYIPIPS--PPRCRIENVTGARCE 47  
 RESULT 6  
 SMDP\_HUMAN STANDARD; PRT; 296 AA.  
 ID 015491;  
 AC 015491;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Neuregulin-1, sensory and motor neuron-derived factor isoform.  
 GN NRG1 OR HGL OR NDF OR HRGA OR GGF OR SMDP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain stem, and Cerebellum;  
 RX MEDLINE=95301541; PubMed=7782315;  
 RA Ho W.-H., Aramant M.P., Nuljens A., Phillips H.S., Osteroff P.L.;  
 RT "Sensory and motor neuron-derived factor. A novel heregulin variant  
 highly expressed in sensory and motor neurons.";  
 RL J. Biol. Chem. 270:14523-14532(1995).  
 CC -1- FUNCTION: THE ISOFORM SMDP MAY PLAY A ROLE IN MOTOR AND SENSORY  
 NEURON DEVELOPMENT.  
 CC -1- SUBCELLULAR LOCATION: SECRETED. MAY POSSESS AN INTERNAL UNCLEAVED  
 SIGNAL SEQUENCE.  
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 10 ISOFORMS OF NRG1 ARE PRODUCED BY  
 ALTERNATIVE SPLICING. EXCEPT FOR SMDP THEY ARE IN ENTRY AC 002297.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN NERVOUS SYSTEM: SPINAL CORD MOTOR  
 NEURONS, DORSAL ROOT GANGLION NEURONS, AND BRAIN. PREDOMINANT  
 ISOFORM EXPRESSED IN SENSORY AND MOTOR NEURONS. NOT DETECTED IN  
 ADULT HEART, PLACENTA, LUNG, LIVER, SKELETAL MUSCLE, KIDNEY, AND  
 PANCREAS. NOT EXPRESSED IN FETAL LUNG, LIVER, AND KIDNEY.  
 CC -1- DEVELOPMENTAL STAGE: HIGHLY EXPRESSED IN DEVELOPING SPINAL MOTOR  
 NEURONS AND IN DEVELOPING CRANIAL NERVE NUCLEI. EXPRESSION IS  
 MAINTAINED ONLY IN BOTH ADULT MOTOR NEURONS AND DORSAL ROOT  
 GANGLION NEURONS.  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: LA1827; AAC41764.1; -  
 DR HSSP: Q12784; 1HRE.  
 DR MIM: 142445; -  
 DR InterPro: IPR000551; EGF-like.  
 DR Pfam: PF00008; EGF\_1.  
 DR SMART: SM00181; EGF\_1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; FALSE\_NEG.  
 KW Growth factor; EGF-like domain; Transmembrane; Multigene family;  
 KM Alternative splicing.  
 FT TRANSMEM 76 100 INTERNAL SIGNAL SEQUENCE (POTENTIAL).  
 FT DOMAIN 58 91 CYS-RICH.  
 FT DOMAIN 211 232 SER/THR-RICH.  
 FT DISULFID 237 251 EGF-LIKE.  
 FT DISULFID 245 265 BY SIMILARITY.  
 FT DISULFID 267 276 BY SIMILARITY.  
 SQ SEQUENCE 296 AA; 31685 MW; 8D41743217F7EB02 CRC64;  
 QY 1 HRPCKDKDLAYCNDGCEVETLTGSHK-CRCEGQGVRCDFL 47  
 DB 233 HLYKCEKEKTEFCVNGEGEFMKDLSNRYICKCPNETTGRCONYV 280  
 Query Match 37.7%; Score 104.5; DB 1; Length 296;  
 Best Local Similarity 31.2%; Pred. No. 1.3e-05;  
 Matches 15; Conservative 14; Mismatches 18; Indels 1; Gaps 1;  
 RESULT 7  
 NR02\_MOUSE STANDARD; PRT; 756 AA.  
 AC P56974;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Pro-neuregulin-2 precursor (Pro-NRG2) [Contains: Neuregulin-2 (NRG-2)  
 DE (Divergent of neuregulin 1) (DON-1)].  
 GN NRG2.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS NRG2-5; NRG2-10 AND NRG2-16A).  
 RC STRAIN-C57BL/6; TISSUE-Brain.  
 RX MEDLINE=9731398; PubMed=918115;  
 RA Carraway K.L., Iit, Weber J.L., Unger M.J., Ledesma J., Yu N.,  
 RA Gasman M., Lai C.;  
 RT "Neuregulin-2, a new ligand of ErbB3/ErbB4-receptor tyrosine  
 RT kinases";  
 RT Nature 387:512-516(1997).  
 RL [2]  
 RP SEQUENCE OF 150-756 FROM N.A. (ISOFORMS DON-1M AND DON-1S).  
 RC TISSUE-Choroid plexus;  
 RX MEDLINE=97342638; PubMed=919335;  
 RA Busfield S.J., Michnick D.A., Chikering T.W., Revett T.L., Ma J.,  
 RA Woolf E.A., Comrack C.A., Dussault B.J., Woolf J., Goodenall A.D.J.,  
 RA Gearing D.P.;  
 RT "Characterization of a neuregulin-related gene, Don-1, that is highly  
 RT expressed in restricted regions of the cerebellum and hippocampus";  
 RL Mol. Cell. Biol. 17:4007-4014(1997).  
 CC -1- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE  
 CC RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORRECEPTORS,

CC RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND  
 CC ACTIVATION OF THE ERBB RECEPTOR. MAY ALSO PROMOTE THE  
 CC HETEROOLIGOMERIZATION WITH THE EGF RECEPTOR.  
 CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS  
 CC A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE  
 CC MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).  
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS: DON-1M, DON-1S/NRG2-5,  
 CC NRG2-10 AND NRG2-16A (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE  
 CC SPLICING.  
 CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN THE BRAIN, WITH LOWER  
 CC LEVELS IN THE LUNG. IN THE CEREBELLUM, FOUND IN GRANULE AND  
 CC PURKINJE CELLS.  
 CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION  
 CC OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE  
 CC PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN  
 CC DIMERIZATION (BY SIMILARITY).  
 CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE  
 CC DOMAIN (BY SIMILARITY).  
 CC -1- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE  
 CC EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR  
 CC FORM (BY SIMILARITY).  
 CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.  
 DR MGD; MGI:1098246; Nrg2.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR003006; IG-MHC.  
 DR InterPro: IPR003598; IG-C2.  
 DR InterPro: IPR002154; Neuregulin.  
 DR Pfam: PF00047; 1g; 1.  
 DR Pfam: PF02158; Neuregulin; 2.  
 DR SMART: SM00181; EGF\_1.  
 DR SMART: SM00408; Igc2; 1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 KW Growth factor; EGF-like domain; Immunoglobulin domain; glycoprotein;  
 KW Transmembrane; Multigene family; Alternative splicing.  
 KM Transmembrane; Multigene family; Alternative splicing.  
 FT PROPEP 1 19  
 FT CHAIN 20 756  
 FT CHAIN 20 314  
 FT CHAIN 20 315  
 FT DOMAIN 20 315  
 FT TRANSMEM 316 336  
 FT DOMAIN 337 756  
 FT DOMAIN 158 226  
 FT DOMAIN 238 248  
 FT DOMAIN 249 290  
 FT DOMAIN 627 633  
 FT DISULFID 165 219  
 FT DISULFID 253 267  
 FT DISULFID 261 278  
 FT DISULFID 280 289  
 FT CARBOHYD 55 55  
 FT CARBOHYD 186 186  
 FT CARBOHYD 254 254  
 FT CARBOHYD 296 296  
 FT CARBOHYD 280 280  
 FT VARSPLIC 281 756  
 FT VARSPLIC 282 330  
 FT VARSPLIC 331 756  
 FT VARSPLIC 282 307  
 FT VARSPLIC 307 756  
 SQ SEQUENCE 756 AA; 82213 MW; 51D5DC91BBE678E CRC64;  
 QY 1 HRPCKDKDLAYCNDGCEVETLTGSHKCRCEGQGVRCDF 46  
 Query Match 37.5%; Score 104; DB 1; Length 756;  
 Best Local Similarity 39.1%; Pred. No. 3.7e-05;  
 Matches 18; Conservative 8; Mismatches 18; Indels 2; Gaps 1;

Db 249 HARCNETAKSYCVNGVCYIEBT--NOLSKCPVGTGRCQOF 292

RESULT 8

NRG1\_RAT STANDARD: PRT: 662 AA.

ID NRG1\_RAT

CC P43322; P43323; P43324; P43325; P43326; P43327; P43328;

CC 16-OCT-2001 (Rel. 40, Created)

CC 16-OCT-2001 (Rel. 40, Last sequence update)

CC 16-OCT-2001 (Rel. 40, Last annotation update)

CC Pro-neuregulin-1 precursor (Pro-NRG) [Contains: Neuregulin-1 (Neu

DE differentiation factor) (Heregulin) (HRG) (Acetylcholine receptor

DE inducing activity) (ARIA) (Sensory and motor neuron-derived factor)

DE (Glia growth factor)].

OS NRG1 OR NDF.

OS Rattus norvegicus (Rat).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

CC NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.

CC TISSUE=Embryoblast;

CC MEDLINE=94158863; PubMed=7509448;

CC Wen D., Suggs S.V., Karunagaran D., Liu N., Cupples R.L., Luo Y.,

CC Janssen A.M., Ben-Baruch N., Trollinger D.B., Jacobsen V.L.,

CC Meng S.-Y., Lu H.S., Hu S., Chang D., Yang W., Yanagihara D.,

CC Koski R.A., Yarden Y.;

CC "Structural and functional aspects of the multiplicity of Neu

RT differentiation factors.";

RL Mol. Cell. Biol. 14:1909-1919(1994).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM ALPHA2C/NDP44), AND PARTIAL SEQUENCE.

CC TISSUE=Embryoblast;

CC MEDLINE=9225796; PubMed=1349853;

CC Wen D., Peles E., Cupples R., Suggs S.V., Bacus S.S., Luo Y.,

CC Trall G., Hu S., Silbiger S.M., Levy R.B., Koski R.A., Lu H.S.,

CC Yarden Y.;

CC "Neu differentiation factor: a transmembrane glycoprotein containing

RT an EGF domain and an immunoglobulin homology unit.";

RL Cell 69:559-572(1992).

RN [3]

RP SEQUENCE OF 14-36.

CC MEDLINE=92208945; PubMed=1348215;

CC Peles E., Bacus S.S., Koski R.A., Lu H.S., Wen D., Ogdén S.G.,

CC Levy R.B., Yarden Y.;

CC "Isolation of the neu/HER-2 stimulatory ligand: a 44 kd glycoprotein

RT that induces differentiation of mammary tumor cells.";

RL Cell 69:205-216(1992).

RN [4]

RP REGULATION OF PROCESSING (ISOFORM ALPHA2C/NDP44).

CC MEDLINE=99069430; PubMed=9852099;

CC Liu X., Hwang H., Cao L., Wen D., Liu N., Graham R.M., Zhou M.;

CC "Release of the neuregulin functional polypeptide requires its

RT cytoplasmic tail.";

RL J Biol. Chem. 273:34335-34340(1998).

RN [5]

RP INTERACTION WITH LIMK1.

CC MEDLINE=98352096; PubMed=9685409;

CC Wang J.Y., Frenzel K.E., Wen D., Falls D.L.;

CC "Transmembrane neuregulins interact with Lim kinase 1, a cytoplasmic

RT protein kinase implicated in development of visuospatial cognition.";

RL J. Biol. Chem. 273:20523-20534(1998).

CC -1- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE

CC RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORCEPTORS,

CC RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND

CC ACTIVATION OF THE ERBB RECEPTORS. THE MULTIPLE ISOFORMS PERFORM

CC DIVERSE FUNCTIONS SUCH AS INDUCING GROWTH AND DIFFERENTIATION OF

CC EPITHELIAL, GLIAL, NEURONAL, AND SKELETAL MUSCLE CELLS; INDUCING

CC EXPRESSION OF ACETYLCHOLINE RECEPTOR IN SYNAPTIC VESICLES DURING

CC THE FORMATION OF THE NEUROMUSCULAR JUNCTION; STIMULATING

CC LOBULOVOLAR BUDDING AND MILK PRODUCTION IN THE MAMMARY GLAND

CC AND INDUCING DIFFERENTIATION OF MAMMARY TUMOR CELLS; STIMULATING

CC SCHWANN CELL PROLIFERATION, IMPLICATION IN THE DEVELOPMENT OF THE

CC MYOCARDIUM SUCH AS TRABECULATION OF THE DEVELOPING HEART (BY

CC SIMILARITY).

CC -1- SUBUNIT: THE CYTOPLASMIC DOMAIN INTERACTS WITH THE LIM DOMAIN

CC REGION OF LIMK1.

CC -1- SUBCELLULAR LOCATION: EXISTS AS A TYPE I MEMBRANE PROTEIN AND AS A

CC PROTOTYPICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE MEMBRANE-

CC BOUND FORM DOES NOT SEEM TO BE ACTIVE.

CC -1- ALTERNATIVE PRODUCTS: AT LEAST 8 ISOFORMS: ALPHA2A/NDP38,

CC ALPHA2B/NDP19, ALPHA2C/NDP44, BETA1, BETA2/NDP40, BETA2A/NDP22,

CC BETA3/NDP4 AND BETA4/NDP42A (SHOWN HERE); ARE PRODUCED BY

CC ALTERNATIVE SPLICING.

CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED. MOST TISSUES CONTAIN ALPHA2A

CC AND ALPHA2B ISOFORMS. ALPHA2 AND BETA2 ARE THE PREDOMINANT FORMS

CC IN MESENCHYMAL AND NONNEURONAL ORGANS. BETA1 IS ENRICHED IN BRAIN

CC AND SPINAL CORD, BUT NOT IN MUSCLE AND HEART. ALPHA2C IS HIGHLY

CC EXPRESSED IN SPINAL CORD, MODERATELY IN LUNG, BRAIN OVARY, AND

CC STOMACH, IN LOW AMOUNTS IN THE KIDNEY, SKIN AND HEART AND NOT

CC DETECTED IN THE LIVER, SPLEEN, AND PLACENTA.

CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION OF

CC TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE

CC PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN

CC DIMERIZATION.

CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE

CC DOMAIN.

CC -1- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE

CC EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR

CC FORM.

CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE.

CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.

CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: U02315; AAA19940.1; -

DR EMBL: U02316; AAA19941.1; -

DR EMBL: U02317; AAA19942.1; -

DR EMBL: U02318; AAA19943.1; -

DR EMBL: U02319; AAA19944.1; -

DR EMBL: U02320; AAA19945.1; -

DR EMBL: U02321; AAA19946.1; -

DR EMBL: U02322; AAA19947.1; -

DR EMBL: U02323; AAA19948.1; -

DR EMBL: U02324; AAA19949.1; -

DR EMBL: W92430; -; NOT\_ANNOTATED\_CDS.

DR HSSP: Q12784; 1HRE.

DR InterPro: IPR000561; EGF-like.

DR InterPro: IPR003006; Ig\_MHC.

DR InterPro: IPR003598; Ig\_C2.

DR InterPro: IPR002154; Neuregulin.

DR Pfam: PF00008; EGF. 1.

DR Pfam: PF00047; Ig. 1.

DR Pfam: PF02158; Neuregulin. 1.

DR PRINTS: PR01089; NEUREGULIN.

DR SMART: SM00181; EGF. 1.

DR SMART: SM00408; IGC2. 1.

DR PROSITE: PS00022; EGF\_1. 1.

DR PROSITE: PS01186; EGF\_2; FALSE\_NEG.

KW Growth factor; EGF-like domain; Immunoglobulin domain; glycoprotein;

KW Transmembrane; Multigene family; Alternative splicing.

KW PROPEP 13

FT CHAIN 14 662 PRO-NEUREGULIN-1, MEMBRANE-BOUND FORM.

FT CHAIN 14 264 NEUREGULIN-1.

FT DOMAIN 14 265 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 266 288 INTERNAL SIGNAL SEQUENCE (POTENTIAL).

FT DOMAIN 289 662 CYTOPLASMIC (POTENTIAL).



FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARSPLIC 1 127 MATSEGLPTSLAPOTDWNSTVPPPLKEMKQEVAV  
 FT GQVLVLCETTSYSPALRFKWLKNGKITKRNENKIK  
 FT KQKYSLEHLYRATLADAGEACRVSKLNDSTKAVIT  
 FT DTNA -> MSEVGETEPSPSAOLSPASISGGPAERNMG  
 FT PHREDSRVGVAGIASCCVCLAEERKGLNSEKICIAI  
 FT LACLLSCICINAKLVYFVDKITEPDSFHLDPGIGODR  
 FT STVDPTLASMVSEVYASFPPIPSLESKREVYQDSLV  
 FT PSRPFLOPISLYNRLIDVGLMSATPSLSPSEPTASQAK  
 FT ATEFNLOTAPKLS (IN ISOFORM BETA1A, ISOFORM  
 FT BETA2A AND ISOFORM BETA2B).  
 FT MISSING (IN ISOFORM BETA2A AND ISOFORM  
 FT BETA2B).  
 FT VARSPLIC 191 198 VSAATTPARNSPVDFHTP -> HTPPTSLLAGKVSILRV  
 FT VARSPLIC 388 405 (IN ISOFORM BETA2B).  
 FT VARSPLIC 406 602 MISSING (IN ISOFORM BETA2B).  
 FT VARSPLIC 602 AA; 67453 MW; 41830856CE5D346 CRC64;  
 SQ SEQUENCE  
 Query Match 35.68; Score 98.5; DB 1; Length 602;  
 Best Local Similarity 33.38; Pred. No. 0.00014;  
 Matches 16; Conservative 11; Mismatches 20; Indels 1; Gaps 1;  
 Oy 1 HFKPCDKDLAYCLNDGCPVITLGSNKH-CRCKEGYGVRCDFL 47  
 Db 137 HLTCKDIKOKAFVNGCECTVWADLPNPRYLCPRCNPFEGRCQNTY 184  
 RESULT 10  
 NR2\_HUMAN STANDARD; PRT; 850 AA.  
 AC 014511;  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Pro-neuregulin-2 precursor (Pro-NRG2) [contains: Neuregulin-2 (NRG-2)  
 DE (Neural and thymus-derived activator for ERBB kinases) (NTAK)  
 DE (Divergent of neuregulin 1) (DON-1)].  
 DE NRG2 OR NTAK.  
 GN Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 CC NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RP TISSUE-Neuroblastoma;  
 RX MEDLINE-98006324; PubMed-9348101;  
 RX Higashiyama S., Horikawa M., Yamada K., Ichino N., Nakano N.,  
 RA Nakagawa T., Miyagawa J., Matsushita N., Nagatsu T., Taniguchi N.,  
 RA Ishiguro H.;  
 RA "A novel brain-derived member of the epidermal growth factor family  
 RT that interacts with ErbB3 and ErbB4.";  
 RL J. Biochem. 122:675-680(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS DON-1B AND DON-1R).  
 RP TISSUE-Fetal brain;  
 RX MEDLINE-97342638; PubMed-9199335;  
 RA Busfield S.J., Michnick D.A., Chickering T.W., Revett T.L., Ma J.,  
 RA Woolf E.A., Comrack C.A., Dussault B.J., Woolf J., Goodenri A.D.J.,  
 RA Gearing D.P.;  
 RA "Characterization of a neuregulin-related gene, Don-1, that is highly  
 RT expressed in restricted regions of the cerebellum and hippocampus.";  
 RL Mol. Cell. Biol. 17:4007-4014(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4; 5 AND 6).  
 RP TISSUE-Lung, and Fetal brain;  
 RX MEDLINE-99295836; PubMed-10369162;  
 RA Ring H.Z., Chang H., Gullbot A., Brice A., Leguern E., Francke U.;  
 RA "The human neuregulin 2 (NRG2) gene: cloning, mapping and evaluation  
 RT as a candidate for the autosomal recessive form of Charcot-Marie-Tooth  
 RT disease linked to 5q.";

RL Hum. Genet. 104:326-332(1999).  
 CC -1- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE  
 CC RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,  
 CC RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND  
 CC ACTIVATION OF THE ERBB RECEPTORS. MAY ALSO PROMOTE THE  
 CC HETERO-DIMERIZATION WITH THE EGF RECEPTOR.  
 CC SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS  
 CC A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE  
 CC MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).  
 CC -1- ALTERNATIVE PRODUCTS: 8 ISOFORMS; 1 (SHOWN HERE), 2, 3, 4, 5, 6,  
 CC DON-1B AND DON-1R. ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: RESTRICTED TO THE CEREBELLUM IN THE ADULT.  
 CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION  
 CC OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE  
 CC PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN  
 CC DIMERIZATION (BY SIMILARITY).  
 CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE  
 CC -1- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE BY THE EGF-LIKE  
 CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY  
 CC FORM (BY SIMILARITY)).  
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 EMBL AB005060. BAA23417.1. -  
 EMBL AF119162. AAF28848.1. -  
 EMBL AF119151. AAF28848.1. JOINED.  
 EMBL AF119152. AAF28848.1. JOINED.  
 EMBL AF119153. AAF28848.1. JOINED.  
 EMBL AF119154. AAF28848.1. JOINED.  
 EMBL AF119155. AAF28848.1. JOINED.  
 EMBL AF119156. AAF28848.1. JOINED.  
 EMBL AF119157. AAF28848.1. JOINED.  
 EMBL AF119158. AAF28848.1. JOINED.  
 EMBL AF119159. AAF28848.1. JOINED.  
 EMBL AF119160. AAF28848.1. JOINED.  
 EMBL AF119161. AAF28848.1. JOINED.  
 EMBL AF119162. AAF28848.1. JOINED.  
 EMBL AF119151. AAF28849.1. JOINED.  
 EMBL AF119152. AAF28849.1. JOINED.  
 EMBL AF119153. AAF28849.1. JOINED.  
 EMBL AF119154. AAF28849.1. JOINED.  
 EMBL AF119155. AAF28849.1. JOINED.  
 EMBL AF119156. AAF28849.1. JOINED.  
 EMBL AF119157. AAF28849.1. JOINED.  
 EMBL AF119158. AAF28849.1. JOINED.  
 EMBL AF119159. AAF28849.1. JOINED.  
 EMBL AF119160. AAF28849.1. JOINED.  
 EMBL AF119161. AAF28849.1. JOINED.  
 EMBL AF119162. AAF28849.1. JOINED.  
 EMBL AF119151. AAF28850.1. JOINED.  
 EMBL AF119152. AAF28850.1. JOINED.  
 EMBL AF119153. AAF28850.1. JOINED.  
 EMBL AF119154. AAF28850.1. JOINED.  
 EMBL AF119155. AAF28850.1. JOINED.  
 EMBL AF119156. AAF28850.1. JOINED.  
 EMBL AF119157. AAF28850.1. JOINED.  
 EMBL AF119158. AAF28850.1. JOINED.  
 EMBL AF119159. AAF28850.1. JOINED.  
 EMBL AF119160. AAF28850.1. JOINED.  
 EMBL AF119161. AAF28850.1. JOINED.  
 EMBL AF119162. AAF28850.1. JOINED.  
 EMBL AF119151. AAF28851.1. JOINED.  
 EMBL AF119152. AAF28851.1. JOINED.  
 EMBL AF119153. AAF28851.1. JOINED.  
 EMBL AF119154. AAF28851.1. JOINED.  
 EMBL AF119155. AAF28851.1. JOINED.  
 EMBL AF119156. AAF28851.1. JOINED.  
 EMBL AF119157. AAF28851.1. JOINED.  
 EMBL AF119158. AAF28851.1. JOINED.  
 EMBL AF119159. AAF28851.1. JOINED.  
 EMBL AF119160. AAF28851.1. JOINED.  
 EMBL AF119161. AAF28851.1. JOINED.  
 EMBL AF119162. AAF28851.1. JOINED.

DR EMBL: AF119157; AAF28851.1; JOINED.  
 DR EMBL: AF119158; AAF28851.1; JOINED.  
 DR EMBL: AF119159; AAF28851.1; JOINED.  
 DR EMBL: AF119160; AAF28851.1; JOINED.  
 DR EMBL: AF119161; AAF28851.1; JOINED.  
 DR EMBL: AF119158; AAF28852.1; JOINED.  
 DR EMBL: AF119151; AAF28852.1; JOINED.  
 DR EMBL: AF119152; AAF28852.1; JOINED.  
 DR EMBL: AF119153; AAF28852.1; JOINED.  
 DR EMBL: AF119154; AAF28852.1; JOINED.  
 DR EMBL: AF119155; AAF28852.1; JOINED.  
 DR EMBL: AF119156; AAF28852.1; JOINED.  
 DR EMBL: AF119157; AAF28853.1; JOINED.  
 DR EMBL: AF119152; AAF28853.1; JOINED.  
 DR EMBL: AF119153; AAF28853.1; JOINED.  
 DR EMBL: AF119154; AAF28853.1; JOINED.  
 DR EMBL: AF119155; AAF28853.1; JOINED.  
 DR EMBL: AF119156; AAF28853.1; JOINED.  
 DR HSSP: 012784; 1HRE.  
 DR MIM: 603818; 1HRE.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR003006; 19-MHC.  
 DR InterPro: IPR003598; 19-C2.  
 DR InterPro: IPR002154; Neuregulin.  
 DR Pfam: PF00008; EGF\_1.  
 DR Pfam: PF00047; 15; 1.  
 DR Pfam: PF02158; Neuregulin; 2.  
 DR SMART: SM00181; EGF\_1.  
 DR SMART: SM00408; IGC2; 1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 DR Growth factor: EGF-like domain; Immunoglobulin domain; Glycoprotein; Transmembrane; Multigene family; Alternative splicing.  
 DR Transmembrane; Multigene family; Alternative splicing.  
 DR PRO-NEUREGULIN-2; MEMBRANE-BOUND FORM.  
 DR CHAIN 112 850  
 DR CHAIN 112 404  
 DR DOMAIN 112 405  
 DR TRANSMEM 406 426  
 DR DOMAIN 427 850  
 DR DOMAIN 250 318  
 DR DOMAIN 340 340  
 DR DOMAIN 341 382  
 DR DOMAIN 10 13  
 DR DOMAIN 20 30  
 DR DOMAIN 33 47  
 DR DOMAIN 87 90  
 DR DOMAIN 721 727  
 DR DISULFID 257 311  
 DR DISULFID 345 359  
 DR DISULFID 353 370  
 DR DISULFID 372 381  
 DR DISULFID 52 52  
 DR CARBOHYD 53 53  
 DR CARBOHYD 147 147  
 DR CARBOHYD 278 278  
 DR CARBOHYD 346 346  
 DR VASPLIC 1 233  
 DR VASPLIC 1 241  
 DR VASPLIC 374 396  
 DR VASPLIC 374 397  
 DR VASPLIC 397 397  
 DR VASPLIC 397 422

FT VARSPLIC 423 850 SSSWSTSPSTLDN (IN ISOFORM 6).  
 FT VARSPLIC 397 426 MISSING (IN ISOFORM 6).  
 FT VARSPLIC 427 850 KAEELYOKRVLTITGICVALLVYICVAV -> SVLMDTP  
 FT VARSPLIC 850 AA: 91678 MW: 7124C089435FD2F4 CRC64;  
 SQ SEQUENCE 850 AA: 91678 MW: 7124C089435FD2F4 CRC64;  
 Query Match 32.5%; Score 90; DB 1; Length 850;  
 Best Local Similarity 34.9%; Pred. No. 0.002;  
 Matches 15; Conservative 9; Mismatches 17; Indels 2; Gaps 1;  
 Oy 1 HRPKPRDCLATCLDGEFVETLTGSHKRCRCGVCVC 43  
 Db 341 HARKNETAKSYCVNGVCYIEGI--NLSCKCPNGFCGRC 381  
 RESULT 11  
 NR02\_RAT STANDARD: PRT: 868 AA.  
 AC 035569; 035570; 035571; 035572; 035073;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Pro-neuregulin-2 precursor (Pro-NRG2) [contains: Neuregulin-2 (NRG-2) (Neutral and thymus-derived activator for ERBB kinases) (NTAK)].  
 DE NRG2 OR NTAK.  
 GN Rattus norvegicus (Rat).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_Taxid-10116;  
 RN [1]  
 RP SEQUENCE FROM N.A., SEQUENCE OF 128-162, AND ALTERNATIVE SPLICING.  
 RX MEDLINE-98006324; PubMed-9348101;  
 RA Higashiyama S., Horikawa M., Yamada K., Ichino N., Nakano N., Nakagawa T., Miyagawa J., Matsushita N., Negatsu T., Taniguchi N., Ishiguro H.;  
 RT "A novel brain-derived member of the epidermal growth factor family that interacts with ErbB3 and ErbB4.";  
 RL J. Biochem. 122:675-680(1997).  
 RN [2]  
 RP SEQUENCE OF 109-868 FROM N.A. (ISOFORMS NRG2-ALPHA AND NRG2-BETA).  
 RX TISSUE=Cerebellum; PubMed-9168114;  
 RA Chang H., Riese D.J. II, Gilbert W., Stern D.F., McMahon U.J.;  
 RT "Ligands for ErbB-family receptors encoded by a neuregulin-like gene.";  
 RL Nature 387:509-512(1997).  
 RL Nature 387:509-512(1997).  
 CC -1- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS, RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND ACTIVATION OF THE ERBB RECEPTORS. MAY ALSO PROMOTE THE HETERODIMERIZATION WITH THE EGF RECEPTOR.  
 CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).  
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 7 ISOFORMS: NTAK-ALPHA (SHOWN HERE), NTAK-ALPHA2, NTAK-ALPHA2B/NTAK-ALPHA2-1P, NTAK-BETA, NTAK-GAMMA, NRG2-ALPHA AND NRG2-BETA. ARE PRODUCED BY ALTERNATIVE SPLICING. THE ALPHA-TYPE AND BETA-TYPE DIFFER IN THE EGF-LIKE DOMAIN.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN MOST PARTS OF THE BRAIN, ESPECIALLY THE OLFACTORY BULB AND CEREBELLUM WHERE IT LOCALIZES IN GRANULE AND PURKINJE CELLS. IN THE HIPPOCAMPUS, FOUND IN THE GRANULE CELLS OF THE DENTATE GYRUS. IN THE BASAL FOREBRAIN, FOUND IN THE MOTOR TRIGEMINAL NUCLEUS. IN THE HINDBRAIN, WEAKLY DETECTABLE IN ADRENAL GLAND, OR TESTIS.  
 CC -1- DEVELOPMENTAL STAGE: IN THE EMBRYO, EXPRESSED IN THE BRAIN OF E11.5 EMBRYOS WHERE IT IS FOUND IN THE TELENEPHALON, BUT NOT IN THE HINDBRAIN. NOT FOUND IN THE HEART. IN THE ADULT, FOUND IN BRAIN AND THYMUS.





FT DOMAIN 145 152 ARG/LYS-RICH (BASIC).  
 FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT DOMAIN 64 104 EGF-LIKE.  
 FT DISULFID 68 81 BY SIMILARITY.  
 FT DISULFID 76 92 BY SIMILARITY.  
 FT DISULFID 94 103 BY SIMILARITY.  
 SQ SEQUENCE 169 AA; 19044 MW; 17F3926ADFB2BDEE CAC64;

Query Match 32.3%; Score 89.5; DB 1; Length 169;  
 Best Local Similarity 40.5%; Pred. No. 0.0005;  
 Matches 17; Conservative 7; Mismatches 15; Indels 3; Gaps 2;

Oy 5 CRDRLAYCNDGECFVETLTGSHKRCCKEGTGVRCDOF 46  
 Db 68 CSSDMNGYCLH-GCCITYLVDN--SONYCRCEGYGVACEHF 106

## RESULT 13

GRA\_SFVKA STANDARD; PRT; 80 AA.  
 AC P08441;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 01-AUG-1988 (Rel. 08, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE Growth factor.  
 OS Shope fibroma virus (strain Kasza) (SFV).  
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Leporipoxvirus.  
 OX NCBI\_TaxID:10272;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-8712751; PubMed-3031480;  
 RA Chang W., Opton C., Hu S.-L., Purchio A.F., McFadden G.;  
 RT "The genome of Shope fibroma virus, a tumorigenic poxvirus, contains  
 a growth factor gene with sequence similarity to those encoding  
 epidermal growth factor and transforming growth factor alpha.";  
 RL Mol. Cell Biol. 7:535-540(1987).  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb-sib.ch](mailto:license@sdb-sib.ch)).  
 CC -----  
 CC EMBL: M15921; AAA66873.1; -  
 DR PIR: A26723; EGVZSF.  
 DR InterPro: IPR000561; EGF-like.  
 DR PRINTS: PRO00136; EGF\_1.  
 DR SMART: SM00181; EGF\_1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; FALSE\_NEG.  
 KW EGF-like domain; growth factor; glycoprotein.  
 FT DOMAIN 29 73  
 FT DISULFID 33 47 EGF-LIKE.  
 FT DISULFID 41 61 BY SIMILARITY.  
 FT DISULFID 63 72 BY SIMILARITY.  
 FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 80 AA; 9210 MW; C48D30E878D2ED58 CRC64;

Query Match 30.7%; Score 85; DB 1; Length 80;  
 Best Local Similarity 39.2%; Pred. No. 0.00087;  
 Matches 20; Conservative 6; Mismatches 17; Indels 8; Gaps 3;

Oy 1 HFKPCRDRLAYCNDGECFVETLTGSHKRCCKEGTGVRCDOF 47  
 Db 29 HAKVCNHDYENYCLNNGCFIADNVSIT--PFCVCRIYNEGSRG-QFI 75

RESULT 14  
 NCCL\_MOUSE STANDARD; PRT; 2531 AA.  
 AC Q01705;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Neurogenic locus notch homolog protein 1 precursor (NOTCH protein).  
 GN NOTCH1 OR NOTCH.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID:10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Embryo;  
 RX MEDLINE-93194170; PubMed-8449489.  
 RA Franco del Amo F., Gendron-Maguire M., Swiatek P.J., Jenkins N.A.,  
 RA Copeland N.G., Gridley T.;  
 RT "Cloning, analysis, and chromosomal localization of Notch-1, a mouse  
 RT homolog of Drosophila Notch.";  
 RL Genomics 15:259-264(1993).  
 RN [2]  
 RP SEQUENCE OF 1551-2170 FROM N.A.  
 RC TISSUE-Embryo;  
 RX MEDLINE-93048835; PubMed-1425352;  
 RA Franco del Amo F., Smith D.E., Swiatek P.J., Gendron-Maguire M.,  
 RA Greenspan R.J., McMahon A.P., Gridley T.;  
 RT "Expression pattern of Notch, a mouse homolog of Drosophila Notch,  
 RT suggests an important role in early postimplantation mouse  
 RT development.";  
 RL Development 115:737-744(1992).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED ALMOST UNIFORMLY IN EARLY EMBRYOS.  
 CC -1- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.  
 CC -1- SIMILARITY: CONTAINS 5 ANK REPEATS.  
 CC -1- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb-sib.ch](mailto:license@sdb-sib.ch)).  
 CC -----  
 CC EMBL: Z11886; CAA77941.1; -  
 DR HSSP: P00740; IEDM.  
 DR MGD: MGI:97363; Notch1.  
 DR InterPro: IPR002110; ANK.  
 DR InterPro: IPR000152; ASX\_hydroxyl.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR000742; EGF\_2.  
 DR InterPro: IPR001881; EGF\_ca.  
 DR InterPro: IPR001438; EGF\_ii.  
 DR InterPro: IPR000800; Notch.  
 DR Pfam: PF00023; ank; 6.  
 DR Pfam: PF00008; EGF; 35.  
 DR Pfam: PF00066; notch; 3.  
 DR PRINTS: PRO0010; EGFBL00D.  
 DR PRINTS: PRO1452; NOTCH.  
 DR SMART: SM00248; ANK; 3.  
 DR SMART: SM00179; EGF\_CA; 23.  
 DR SMART: SM00001; EGF\_LIKE; 11.  
 DR SMART: SM00004; NL; 2.  
 DR PROSITE: PS50088; ANK\_REPEAT; 2.  
 DR PROSITE: PS50297; ANK\_REPEAT\_REGION; 1.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; 22.  
 DR PROSITE: PS00022; EGF\_1; 34.  
 DR PROSITE: PS01186; EGF\_2; 27.



Query Match	Best Local Similarity	Score	Length
Db 1064	DSAPCKNGRCW----	QTNTQYHCECRSGWTGVNCD	1095
Matches	15;	Conservative	4;
Mismatches	13;	Indels	4;
Gaps	1;		

DT 01-FEB-1994 (Rel. 28, last sequence update)  
 DT 16-OCT-2001 (Rel. 40, last annotation update)  
 DE Betacellulin precursor (BTC).  
 GN BTC OR BCN.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 32-54; 64-71 AND 75-111.  
 RC TISSUE-Pancreas;  
 RX MEDLINE=93206093; PubMed=8456283;  
 RA Shing Y., Christofori G., Hanahan D., Ono Y., Sasada R.,  
 RA Igarashi K., Folkman J.;  
 RT "Betacellulin: a mitogen from pancreatic beta cell tumors.";  
 RL Science 259:1604-1607 (1993).  
 CC -1- FUNCTION: POTENT MITOGEN FOR RETINAL PIGMENT EPITHELIAL CELLS  
 CC AND VASCULAR SMOOTH MUSCLE CELLS. THE EFFECTS OF BETACELLULIN  
 CC ARE PROBABLY MEDIATED BY THE EGF RECEPTOR AND OTHER RELATED  
 CC RECEPTORS.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (PRECURSOR FORM);  
 CC EXTRACELLULAR (MATURE FORM).  
 CC -1- TISSUE SPECIFICITY: FOUND IN SEVERAL MOUSE TISSUES INCLUDING  
 CC KIDNEY, UTERUS AND LIVER AS WELL AS IN BETA TUMOR CELL LINE AND  
 CC MCF-7 CELLS. IT IS NOT DETECTED IN THE BRAIN.  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L08394; AAA40511.1; -  
 DR PIR; A37408; A37408.  
 DR HSSP; P01135; 3TGF.  
 DR MGD; MGI:99439; Btc.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR01336; EGF\_1.  
 DR Pfam; PF00008; EGF; 1.  
 DR PRINTS; PRO0009; EGFTEF.  
 DR SMART; SM00181; EGF; 1.  
 DR SMART; SM00001; EGF-like; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR Growth factor; Mitogen; Glycoprotein; EGF-like domain; Transmembrane;  
 KW signal.  
 FT SIGNAL. 1 31  
 FT CHAIN 32 111  
 FT PROPEP 112 177  
 FT DOMAIN 32 118  
 FT TRANSMEM 119 139  
 FT DOMAIN 140 177  
 FT DOMAIN 65 105  
 FT DOMAIN 146 153  
 FT DISULFID 69 82  
 FT DISULFID 77 93  
 FT DISULFID 95 104  
 FT CARBOHYD 34 34  
 FT CARBOHYD 42 42  
 FT CARBOHYD 52 52  
 SO SEQUENCE 177 AA; 19664 MW; 06BB34F0E13F82B CRC64;

Query Match 29.8%; Score 82.5; DB 1; Length 177;  
 Best Local Similarity 39.1%; Pred. No. 0.0036;  
 Matches 18; Conservative 9; Mismatches 14; Indels 5; Gaps 3;

QY 1 HFKPCKDYLAVCLINDSEC-FVETITLGSKHCKRGYGVRCQ 45  
 DB 65 HFSRCPRQYKHNCIH-GRCRFVVDQTPS---CICERGFYGCRCR 106

Search completed: September 13, 2002, 08:55:43  
 Job time: 350 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 08:44:43 ; Search time 18.37 Seconds  
(without alignments)  
245.846 Million cell updates/sec

Title: US-09-877-665-4

Perfect score: 277  
Sequence: 1 HFKPCRDKLAYCLNDGECPYETLTGSHKRCCKEGYGVRCDOFL 47

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 71: \*  
1: PIR1: \*  
2: PIR2: \*  
3: PIR3: \*  
4: PIR4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	277	100.0	713	2	T44447
2	113.5	41.0	125	2	I38405
3	113.5	41.0	462	2	I38405
4	113.5	41.0	640	2	A43273
5	111.5	40.3	639	2	I61719
6	103.5	38.1	125	2	S62676
7	104.5	37.7	175	2	I38408
8	104.5	37.7	241	2	S32359
9	104.5	37.7	241	2	D43273
10	104.5	37.7	296	2	A56943
11	104.5	37.7	422	2	S32357
12	104.5	37.7	637	2	C43273
13	104.5	37.7	645	2	B43273
14	102.5	37.0	230	2	A56210
15	102.5	37.0	636	2	I61718
16	102.5	37.0	662	2	I61722
17	98.5	35.6	602	2	A45769
18	92.5	33.4	2180	2	T29764
19	90	32.5	850	2	JC5700
20	90	32.5	860	2	JC5702
21	90	32.5	868	2	JC5701
22	85	30.7	80	1	ECVZSF
23	84.5	30.5	1220	2	A56136
24	83	30.0	46	2	JT0747
25	83	30.0	162	2	S68401
26	83	30.0	861	2	A48825
27	83	30.0	2531	2	A46019
28	82.5	29.8	177	2	A37408
29	82	29.6	85	1	EGVZM1

30	82	29.6	230	2	A44074	probable EGF-like
31	82	29.6	264	2	T22380	hypothetical prote
32	80.5	29.1	178	2	JC1467	betacellulin precu
33	80	28.9	907	2	T27317	hypothetical prote
34	79.5	28.7	140	1	WMV29	growth factor - va
35	79.5	28.7	140	2	T30766	growth factor - va
36	79.5	28.7	142	1	WMV23C	transforming growt
37	79.5	28.7	159	1	WFR11	transforming growt
38	79.5	28.7	159	1	S27195	transforming growt
39	79.5	28.7	159	2	I57497	transforming growt
40	79	28.5	2531	2	S18188	transforming growt
41	79	28.5	1207	1	EGH0	transforming growt
42	77.5	28.0	722	2	I48324	transforming growt
43	77.5	28.0	2352	2	T30201	transforming growt
44	76.5	27.6	1372	2	T25933	transforming growt
45	76	27.4	482	2	JC5092	transforming growt

## ALIGNMENTS

RESULT 1  
T44447  
neuregulin-3 [imported] - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jul-2000  
R:Zhang, D.; Sliwkowski, M.X.; Mart, M.; Frantz, G.; Akita, R.; Sun, Y.; Hillan, K.;  
Proc. Natl. Acad. Sci. U.S.A. 94, 9562-9567, 1997  
A:Title: Neuregulin-3 (NRG3): A novel neural tissue-enriched protein that binds and a  
A:Reference number: 222773; MUID:97420720  
A:Accession: T44447  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-713 <ZHA>  
A:Cross-references: EMBL:AF010130; NID:92429163; PIDN:AA870914.1; PID:92429164  
C:Genetics:  
A:Gene: NRG3

Query Match 100.0%; Score 277; DB 2; Length 713;  
Best Local Similarity 100.0%; Pred. No. 4.7e-24;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HFKPCRDKLAYCLNDGECPYETLTGSHKRCCKEGYGVRCDOFL 47  
DB 288 HFKPCRDKLAYCLNDGECPYETLTGSHKRCCKEGYGVRCDOFL 334

RESULT 2  
I38405  
neu differentiation factor - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 05-Nov-1999  
R:Men, D.; Suggs, S.V.; Karunakaran, D.; Lau, N.; Cupples, R.L.; Luo, Y.; Janssen, A.  
Mol. Cell. Biol. 14, 1909-1919, 1994  
A:Title: Structural and functional aspects of the multiplicity of Neu differentiation  
A:Reference number: A56210; MUID:94158863  
A:Accession: I38405  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-125 <RES>  
A:Cross-references: EMBL:U02327; NID:9408404; PIDN:AAA19952.1; PID:9408405

Query Match 41.0%; Score 113.5; DB 2; Length 125;  
Best Local Similarity 34.8%; Pred. No. 5.1e-06;  
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;  
OY 1 HFKPCRDKLAYCLNDGECPYETLTGSHKRCCKEGYGVRCDO 45  
DB 56 HFKPCRDKLAYCLNDGECPYETLTGSHKRCCKEGYGVRCDO 101

```

RESULT      3
138404
neu differentiation factor - human
C:Species: Homo sapiens (man)
C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 05-Nov-1999
C:Accession: 138404
R:Wen, D.; Suggs, S.V.; Karunagaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.M.;
Mol. Cell. Biol. 14, 1909-1919, 1994
A:Title: Structural and functional aspects of the multiplicity of Neu differentiation fa
A:Reference number: A56210; MUID:94158863
A:Accession: 138404
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-462 <RES>
A:Cross-references: EMBL:U02326; NID:9408402; PIDN:AAA19951.1; PID:9408403

```

```

Query Match      41.0%; Score 113.5; DB 2; Length 462;
Best Local Similarity 34.8%; Pred. No. 1.6e-05;
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

OY      1 HEKPCRDKLAVCLNDGECFVETLTGSHKH-CRCKEGYGVRCDO 45
DB      178 HLKCAEKETFCVNGGECFVWDLNPSRYLCKPCPGFGARCTE 223

```

```

RESULT      4
A43273
heregulin precursor, splice form alpha - human
N:Alternate names: breast cancer cell differentiation factor p45; Neu differentiation fa
C:Species: Homo sapiens (man)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 11-Jan-2000
C:Accession: A43273; A48498; A38155
R:Holmes, W.E.; Sliwowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Yansut
Science 259, 1205-1210, 1992
A:Title: Identification of heregulin, a specific activator of p185(erbB2).
A:Reference number: A43273; MUID:92271253
A:Accession: A43273
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-640 <HOL>
A:Experimental source: breast tumor cell line, MDA-MB-231, ATCC HTB 26
A:Note: sequence extracted from NCBI backbone (NCBIP:103250)
R:Culioscu, J.M.; Plozman, G.D.; Carlton, G.W.; Green, J.M.; Shoyab, M.
J. Biol. Chem. 268, 18407-18410, 1993
A:Title: Characterization of a breast cancer cell differentiation factor that specifies
A:Reference number: A48498; MUID:93366731
A:Accession: A48498
A:Molecule type: protein
A:Residues: 20-21, 'X', '23-24', 'XX', '27-28 <CUU>
R:Peles, E.; Bacus, S.S.; Koshi, R.A.; Lu, H.S.; Wen, D.; Ogden, S.G.; Levy, R.B.; Yarde
Cell 69, 205-216, 1992
A:Title: Isolation of the neu/HER-2 stimulatory ligand: a 44 kd glycoprotein that induce
A:Reference number: A38155; MUID:92208945
A:Accession: A38155
A:Molecule type: protein
A:Residues: 'X', '15-16', 'X', '18-20', 'RG', '23-24', 'GP', '27', 'E', '29', 'XP', '32-36 <PEL>
A:Note: sequence extracted from NCBI backbone (NCBIP:91347)
C:Genetics:
A:Gene: GDB:HGL
A:Cross-references: GDB:132656; OMIM:142445
A:Map position: 8p22-8p11
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: alternative splicing; glycoprotein
F:182-221/Domain: EGF homology <EGF>

```

```

Query Match      41.0%; Score 113.5; DB 2; Length 640;
Best Local Similarity 34.8%; Pred. No. 2.1e-05;
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

```

```

OY      1 HEKPCRDKLAVCLNDGECFVETLTGSHKH-CRCKEGYGVRCDO 45
DB      178 HLKCAEKETFCVNGGECFVWDLNPSRYLCKPCPGFGARCTE 223

```

```

RESULT      5
161719
neu differentiation factor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 01-Dec-2000
C:Accession: 161719; 161723; 161716; 161717; 161724; A38220
R:Wen, D.; Suggs, S.V.; Karunagaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A
Mol. Cell. Biol. 14, 1909-1919, 1994
A:Title: Structural and functional aspects of the multiplicity of Neu differentiatio
A:Reference number: A56210; MUID:94158863
A:Accession: 161719
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-639 <RES>
A:Cross-references: EMBL:U02319; NID:9408388; PIDN:AAA19944.1; PID:9408389
A:Accession: 161723
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-639 <RES>
A:Cross-references: EMBL:U02323; NID:9408396; PIDN:AAA19948.1; PID:9408397
A:Accession: 161716
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-422, 'H', 'NL', '637-638', 'ELRNKAYRSKCMQIOLSATHLRPSITHLGFTL' <RE3>
A:Cross-references: EMBL:U02316; NID:9408384; PIDN:AAA19942.1; PID:9408385
A:Accession: 161717
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-422, 'H', 'NL', '637-638', 'ELRNKAYRSKCMQIOLSATHLRPSITHLGFTL' <RE4>
A:Cross-references: EMBL:U02317; NID:9408384; PIDN:AAA19942.1; PID:9408385
A:Accession: 161724
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-422 <RES>
A:Cross-references: EMBL:U02324; NID:9408398; PIDN:AAA19949.1; PID:9408399
R:Wen, D.; Peles, E.; Cupples, R.; Suggs, S.V.; Bacus, S.S.; Luo, Y.; Trall, G.; Hu,
Cell 69, 559-572, 1992
A:Title: Neu differentiation factor: a transmembrane glycoprotein containing an EGF
A:Reference number: A38220; MUID:92257556
A:Accession: A38220
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-422 <WEN>
A:Note: sequence extracted from NCBI backbone (NCBIN:101767, NCBIP:101768)
C:Superfamily: unassigned EGF-related proteins; EGF homology

```

```

Query Match      40.3%; Score 111.5; DB 2; Length 639;
Best Local Similarity 34.8%; Pred. No. 3.5e-05;
Matches 16; Conservative 13; Mismatches 16; Indels 1; Gaps 1;

OY      1 HEKPCRDKLAVCLNDGECFVETLTGSHKH-CRCKEGYGVRCDO 45
DB      178 HLKCAEKETFCVNGGECFVWDLNPSRYLCKPCPGFGARCTE 223

```

```

RESULT      6
S62676
heregulin isoform alpha 2 - human (fragments)
N:Alternate names: differentiation factor neu isoform alpha 2
C:Species: Homo sapiens (man)
C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 13-Mar-1997
C:Accession: S62676
R:Hara, S.; Liu, N.; Meng, S.Y.; Lu, H.S.
Biochim. Biophys. Acta 1292, 168-176, 1996
A:Title: Isolation and structural characterization of recombinant human neu different
A:Reference number: S62676; MUID:96139341
A:Accession: S62676

```

A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-6;7-16;17-30;31-38;39-58;59-92;93-120;121-125 <NAR>  
C:Keywords: proto-oncogene

Query Match 38.1%; Score 105.5; DB 2; Length 125;  
Best Local Similarity 32.6%; Pred. No. 4.2e-05;  
Matches 15; Conservative 14; Mismatches 16; Indels 1; Gaps 1;

OY 1 HFKPRDKDLAYCLNDGECFVETLTGSHKH-CRCKEGYGVRCDO 45  
DB 75 HLVKCAEKERTFCVNGGECFVWKDLSNPSRYLCKPCPNETGTDRCONY 120

RESULT 7  
I38408  
neu differentiation factor - human (fragment)  
C:Species: Homo sapiens (man)  
C>Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 11-Jan-2000  
C:Accession: I38408  
R:Men, D.; Suggs, S.V.; Katunaga, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.M.; Mol. Cell. Biol. 14, 1909-1919, 1994  
A>Title: Structural and functional aspects of the multiplicity of Neu differentiation factor  
A:Reference number: A56210; MUID:94158863  
A:Accession: I38408  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-175 <RES>  
A:Cross-references: EMBL:U02330; NID:9408410; PIDN:AAA19955.1; PID:9408411  
C:Superfamily: unassigned EGF-related proteins; EGF homology  
F:116-135/Domain: EGF homology <EGF>

Query Match 37.7%; Score 104.5; DB 2; Length 175;  
Best Local Similarity 31.2%; Pred. No. 7.3e-05;  
Matches 15; Conservative 14; Mismatches 18; Indels 1; Gaps 1;

OY 1 HFKPRDKDLAYCLNDGECFVETLTGSHKH-CRCKEGYGVRCDO 47  
DB 112 HLVKCAEKERTFCVNGGECFVWKDLSNPSRYLCKPCPNETGTDRCONY 159

RESULT 8  
S32359  
glial growth factor - bovine  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 19-Mar-1997 #sequence\_revision 01-Aug-1997 #text\_change 11-Jan-2000  
C:Accession: S32359  
R:Marchionni, M.A.; Goodearl, A.D.J.; Chen, M.S.; Bermingham-McDonogh, O.; Kirk, C.; Herles, I.; Davis, J.B.; Hsuan, J.J.; Totty, N.F.; Otsu, M.; McBurney, R.N.; Waterfield, M. Nature 362, 312-318, 1993  
A>Title: Glial growth factors are alternatively spliced erbB2 ligands expressed in the brain  
A:Reference number: S32357; MUID:93205115  
A:Accession: S32359  
A:Molecule type: mRNA  
A:Residues: 1-241 <MAR>  
A:Cross-references: GB:U2259; NID:9289413; PIDN:AAA30540.1; PID:9289414  
C:Superfamily: unassigned EGF-related proteins; EGF homology  
F:182-221/Domain: EGF homology <EGF>

Query Match 37.7%; Score 104.5; DB 2; Length 241;  
Best Local Similarity 31.2%; Pred. No. 9.6e-05;  
Matches 15; Conservative 14; Mismatches 18; Indels 1; Gaps 1;

OY 1 HFKPRDKDLAYCLNDGECFVETLTGSHKH-CRCKEGYGVRCDO 47  
DB 178 HLVKCAEKERTFCVNGGECFVWKDLSNPSRYLCKPCPNETGTDRCONY 225

RESULT 9

D43273  
heregulin precursor, splice form beta-3 - human  
N:Alternate names: glial growth factor HRC-beta-3; heregulin  
C:Species: Homo sapiens (man)  
C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 11-Jan-2000  
C:Accession: D43273; S32358  
R:Holmes, W.E.; Sliwkowski, M.X.; Akita, R.M.; Hensel, W.J.; Lee, J.; Park, J.W.; Yan Science 256, 1205-1210, 1992  
A>Title: Identification of heregulin, a specific activator of p185(erbB2).  
A:Reference number: A43273; MUID:92271253  
A:Accession: D43273  
A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual  
A:Molecule type: mRNA  
A:Residues: 1-241 <HOL>  
R:Marchionni, M.A.; Goodearl, A.D.J.; Chen, M.S.; Bermingham-McDonogh, O.; Kirk, C.; Herles, I.; Davis, J.B.; Hsuan, J.J.; Totty, N.F.; Otsu, M.; McBurney, R.N.; Waterfield, M. Nature 362, 312-318, 1993  
A>Title: Glial growth factors are alternatively spliced erbB2 ligands expressed in the brain  
A:Reference number: S32357; MUID:93205115  
A:Accession: S32358  
A:Molecule type: mRNA  
A:Residues: 1-241 <MAR>  
A:Cross-references: GB:U2261; NID:9292049; PIDN:AAB59358.1; PID:9292050  
C:Genetics:  
A:Gene: GDB:HGL; GGF  
A:Cross-references: GDB:132656; OMIM:142445  
A:Map position: 8p22-8p11  
C:Superfamily: unassigned EGF-related proteins; EGF homology  
C:Keywords: alternative splicing  
F:182-221/Domain: EGF homology <EGF>

Query Match 37.7%; Score 104.5; DB 2; Length 241;  
Best Local Similarity 31.2%; Pred. No. 9.6e-05;  
Matches 15; Conservative 14; Mismatches 18; Indels 1; Gaps 1;

OY 1 HFKPRDKDLAYCLNDGECFVETLTGSHKH-CRCKEGYGVRCDO 47  
DB 178 HLVKCAEKERTFCVNGGECFVWKDLSNPSRYLCKPCPNETGTDRCONY 225

RESULT 10  
A56943  
sensory/motor neuron-derived factor - human  
C:Species: Homo sapiens (man)  
C>Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 11-Jan-2000  
C:Accession: A56943  
R:Ho, W.H.; Armanini, M.P.; Nuljens, A.; Phillips, H.S.; Osheroff, P.L. J. Biol. Chem. 270, 14523-14532, 1995  
A>Title: Sensory and motor neuron-derived factor. A novel heregulin variant highly expressed in the brain  
A:Reference number: A56943; MUID:95301541  
A:Accession: A56943  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-296 <HOL>  
A:Cross-references: GB:U41827; NID:9862422; PIDN:AAC41764.1; PID:9862423  
C:Superfamily: unassigned EGF-related proteins; EGF homology  
F:237-276/Domain: EGF homology <EGF>

Query Match 37.7%; Score 104.5; DB 2; Length 296;  
Best Local Similarity 31.2%; Pred. No. 0.00011;  
Matches 15; Conservative 14; Mismatches 18; Indels 1; Gaps 1;

OY 1 HFKPRDKDLAYCLNDGECFVETLTGSHKH-CRCKEGYGVRCDO 47  
DB 233 HLVKCAEKERTFCVNGGECFVWKDLSNPSRYLCKPCPNETGTDRCONY 280

RESULT 11  
S32357  
glial growth factor - human  
C:Species: Homo sapiens (man)









PT Infection, malignancy, Alzheimer's disease or Down's syndrome  
XX  
XX Claim 30; Page 64; 101pp; English.

CC This is the epidermal growth factor (EGF)-like domain of human  
CC neuroligin related ligand NR33 (see also AAW97618), a novel member of  
CC the EGF-like family of protein ligands that binds to the ErbB4  
CC receptor and activates ErbB4 receptor tyrosine phosphorylation.  
CC The EGF-like domain of NR33 is distinct from the EGF-like domains  
CC of NR31 and NR32. The invention provides human and murine  
CC polypeptides (see also AAW97617) that have at least 75% homology to  
CC the NR33 EGF-like domain, as well as expression vectors, host cells  
CC and methods for the recombinant production of novel NR33s. The  
CC NR33 polypeptides and polynucleotides and can be used to enhance  
CC the survival, proliferation or differentiation of cells having the  
CC ErbB4 receptor in vivo and in vitro. They can be used to prevent  
CC or treat damage to a nerve or damage to other NR33-expressing or  
CC NR33-responsive cells, e.g. brain, heart, or kidney cells. In  
CC particular, they can be used to treat diseases which involve neural  
CC cell growth such as demyelination, or damage or loss of glial cells  
CC (e.g. multiple sclerosis). They can be used to treat patients whose  
CC nervous system has been damaged by e.g. trauma, surgery, stroke,  
CC ischemia, infection, metabolic disease, nutritional deficiency,  
CC malignancy, or toxic agents. NR33 can also be used to treat  
CC motor neuron disorders such as amyotrophic lateral sclerosis (Lou  
CC Gehrig's disease), Bell's palsy, conditions involving spinal  
CC muscular atrophy or paralysis, neurodegenerative disorders such as  
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple  
CC sclerosis, Huntington's chorea, Down's syndrome, nerve deafness,  
CC and Meniere's disease. They can also be used to treat neuropathies  
CC associated with systemic disease including post-polio syndrome,  
CC hereditary neuropathies including Charcot-Marie-Tooth disease,  
CC Refsum's disease, abetalipoproteinemia, Tangier disease, Krabbe's  
CC disease, metachromatic leukodystrophy, Fabry's disease and  
CC Dejerine-Sottas syndrome, to treat disease of skeletal muscle of  
CC smooth muscle, such as muscular dystrophy or diseases caused by  
CC skeletal or smooth muscle wasting. The products can also be used  
CC for detection, diagnosis, for the production of transgenic or  
CC knockout animals or for drug screening. A claimed immunoglobulin  
CC sequence.

XX Sequence 47 AA;

QY Query Match 100.0%; Score 277; DB 20; Length 47;  
Best Local Similarity 100.0%; Pred. No. 2.5e-21;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 HFKPCRDRLAVCLNDGECFVETLTGSHKRCRCEGYGVRCDFL 47  
1 hfkpcrdkrlavclndgectfvetltgshkrcrcegygvrcdfll 47

RESULT 2

AA05451 ID AAY05451 standard; Protein; 157 AA.

XX AC AAY05451;

DT 06-JUL-1999 (first entry)

DE Human heregulin-like factor sequence.

KW Human heregulin-like factor; HLF; cell growth regulator; diagnosis;  
neural system disorder; cancer.

OS Homo sapiens.

XX MO9857989-A1.

PD 23-DEC-1998.

*X not good dots*

PF 16-JUN-1998; 98MO-US12403.

XX 17-JUN-1997; 97OS-0049942.

PA (HUMA-) HUMAN GENOME SCI INC.

XX (GEU) UNIV GEORGETOWN.

PI H1az1 MM, King CR, Ruben SM, Young P;

XX WPI; 1999-095327/08.

DR N-PSDB; AAX36423.

PT New isolated heregulin-like factor - used to develop products for  
PT the diagnosis and treatment of disorders involving regulation of  
PT cell growth, particularly cancers

PS Claim 17; Page 86-87; 118pp; English.

CC This sequence is the human heregulin-like factor (HLF) of the  
CC invention. The HLF is involved in the regulation of cell growth.  
CC Detection of different levels of expression of the HLF gene can be used  
CC for the diagnosis of disorders, e.g. in the neural system. In  
CC particular, detection of different levels of HLF gene expression in cells  
CC or body fluid of an individual can be used for diagnosing cancer. The  
CC products can also be used in the treatment of disorders involving  
CC abnormal levels of HLF activity.

XX Sequence 157 AA;

QY Query Match 100.0%; Score 277; DB 20; Length 157;  
Best Local Similarity 100.0%; Pred. No. 7.6e-21;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 HFKPCRDRLAVCLNDGECFVETLTGSHKRCRCEGYGVRCDFL 47  
31 hfkpcrdkrlavclndgectfvetltgshkrcrcegygvrcdfll 77

RESULT 3

AA097621 ID AAW97621 standard; Protein; 360 AA.

XX AC AAW97621;

DT 10-MAY-1999 (first entry)

DE Human neuroligin related ligand NR33 extracellular domain.

KW Neuroligin related ligand; NR33; hNR33l; human; ErbB4 receptor;  
signal transduction; nervous system disorder; neurodegeneration;  
neuropathy; therapy; diagnosis.

OS Homo sapiens.

XX MO9902681-A1.

PD 21-JAN-1999.

DE 30-JUN-1998; 98MO-US13411.

PR 24-JUL-1997; 97OS-0899437

PR 09-JUL-1997; 97OS-0052049.

PA (GETH) GENENTECH INC.

XX Godowski P, Mark MR, Zhang D;

XX WPI; 1999-120882/10.

PT New isolated neuroligin related ligand-3 - used to develop products  
PT for treating nervous system disorders, e.g. stroke, ischemia,  
PT infection, malignancy, Alzheimer's disease or Down's syndrome

*X not good dots*

*98MO*

PS Claim 5(a); Page 69-70; 101pp; English.

XX This is the extracellular domain (ECD, aa1-360 of human neuregulin  
XX related ligand NRG3 (see also AAM97618), a novel member of the  
CC epidermal growth factor (EGF)-like family of protein ligands. NRG3  
CC binds to the ErbB4 receptor, but not to the ErbB2 or ErbB3 receptor,  
CC activates ErbB4 receptor tyrosine phosphorylation. The invention  
CC provides human and murine polypeptides (see also AAM97617) that have  
CC at least 75% homology to the NRG3 ECD, as well as expression vectors,  
CC host cells and methods for the recombinant production of novel  
CC NRG3s. The NRG3 polypeptides and polynucleotides and can be used to  
CC enhance the survival, proliferation or differentiation of cells  
CC having the ErbB4 receptor in vivo and in vitro. They can be used to  
CC prevent or treat damage to a nerve or damage to other NRG3-expressing  
CC or NRG3-responsive cells, e.g. brain, heart, or kidney cells. In  
CC particular, they can be used to treat diseases which involve neural  
CC cell growth such as demyelination, or damage or loss of glial cells  
CC (e.g. multiple sclerosis). They can be used to treat patients whose  
CC nervous system has been damaged by e.g. trauma, surgery, stroke,  
CC ischemia, infection, metabolic disease, nutritional deficiency,  
CC malignancy, or toxic agents. NRG3 can also be used to treat  
CC motor neuron disorders such as amyotrophic lateral sclerosis (Lou  
CC Gehrig's disease), Bell's palsy, conditions involving spinal  
CC muscular atrophy or paralysis, neurodegenerative disorders such as  
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple  
CC sclerosis, Huntington's chorea, Down's syndrome, nerve deafness,  
CC and Meniere's disease. They can also be used to treat neuropathies  
CC associated with systemic disease including post-polio syndrome,  
CC hereditary neuropathies including Charcot-Marie-Tooth disease,  
CC Refsum's disease, abetalipoproteinemia, Tange's disease, Krabbe's  
CC disease, metachromatic leukodystrophy, Fabry's disease and  
CC Dejerine-Sottas syndrome, to treat disease of skeletal muscle of  
CC skeletal or smooth muscle wasting. The products can also be used  
CC for detection, diagnosis, for the production of transgenic or  
CC knockout animals or for drug screening.

Sequence 360 AA;

Query Match 100.0%; Score 277; DB 20; Length 360;  
Best Local Similarity 100.0%; Pred. No. 1.6e-20;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFRPCRDKLAYCLNDGEVFIEITLGSNKRCKRCKEGYGVRCDOFL 47  
|||  
DB 286 hfkpcrdklayclndgecvletltgshkrcckrckegygvcrcdql 332

RESULT 4

ID AAM97620 standard; Protein; 362 AA.

AAW97620;

10-MAY-1999 (first entry)

Mouse neuregulin related ligand NRG3 extracellular domain.

Neuregulin related ligand; NRG3; mouse; ErbB4 receptor;  
signal transduction; nervous system disorder; neurodegeneration;  
neuropathy; therapy; diagnosis.

Mus sp.

WO9902681-A1.

21-JAN-1999.

30-JUN-1998; 98WO-US13411.

24-JUL-1997; 97US-0899437.

PR 09-JUL-1997; 97US-0052019.  
XX GENENTECH INC.  
XX Godowski PJ  
XX Maix MR, Zhang D;  
XX Ref. 1999-120882/10.

PS New isolated neuregulin related ligand-3 - used to develop products  
PT for treating nervous system disorders, e.g. stroke, ischemia,  
PT infection, malignancy, Alzheimer's disease or Down's syndrome

Claim 5(a); Page 62-63; 101pp; English.

XX This is the extracellular domain (ECD, aa1-362) of murine neuregulin  
XX related ligand NRG3 (see also AAM97617), a novel member of the  
CC epidermal growth factor (EGF)-like family of protein ligands. NRG3  
CC binds to the ErbB4 receptor, but not to the ErbB2 or ErbB3 receptor,  
CC activates ErbB4 receptor tyrosine phosphorylation. The invention  
CC provides human and murine polypeptides (see also AAM97618) that have  
CC at least 75% homology to the NRG3 ECD, as well as expression vectors,  
CC host cells and methods for the recombinant production of novel  
CC NRG3s. The NRG3 polypeptides and polynucleotides and can be used to  
CC enhance the survival, proliferation or differentiation of cells  
CC having the ErbB4 receptor in vivo and in vitro. They can be used to  
CC prevent or treat damage to a nerve or damage to other NRG3-expressing  
CC or NRG3-responsive cells, e.g. brain, heart, or kidney cells. In  
CC particular, they can be used to treat diseases which involve neural  
CC cell growth such as demyelination, or damage or loss of glial cells  
CC (e.g. multiple sclerosis). They can be used to treat patients whose  
CC nervous system has been damaged by e.g. trauma, surgery, stroke,  
CC ischemia, infection, metabolic disease, nutritional deficiency,  
CC malignancy, or toxic agents. NRG3 can also be used to treat  
CC motor neuron disorders such as amyotrophic lateral sclerosis (Lou  
CC Gehrig's disease), Bell's palsy, conditions involving spinal  
CC muscular atrophy or paralysis, neurodegenerative disorders such as  
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple  
CC sclerosis, Huntington's chorea, Down's syndrome, nerve deafness,  
CC and Meniere's disease. They can also be used to treat neuropathies  
CC associated with systemic disease including post-polio syndrome,  
CC hereditary neuropathies including Charcot-Marie-Tooth disease,  
CC Refsum's disease, abetalipoproteinemia, Tange's disease, Krabbe's  
CC disease, metachromatic leukodystrophy, Fabry's disease and  
CC Dejerine-Sottas syndrome, to treat disease of skeletal muscle of  
CC skeletal or smooth muscle wasting. The products can also be used  
CC for detection, diagnosis, for the production of transgenic or  
CC knockout animals or for drug screening.

Sequence 362 AA;

Query Match 100.0%; Score 277; DB 20; Length 362;  
Best Local Similarity 100.0%; Pred. No. 1.6e-20;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HFRPCRDKLAYCLNDGEVFIEITLGSNKRCKRCKEGYGVRCDOFL 47  
|||  
DB 288 hfkpcrdklayclndgecvletltgshkrcckrckegygvcrcdql 334

RESULT 5

ID AAM97619 standard; Protein; 696 AA.

AAW97619;

10-MAY-1999 (first entry)

Human neuregulin related ligand NRG3 (splice variant).

Neuregulin related ligand; NRG3; human; ErbB4 receptor;  
signal transduction; nervous system disorder; neurodegeneration;

neuropathy; therapy; diagnosis; splice variant.  
 Homo sapiens.  
 Key Location/Qualifiers  
 1..360 "extracellular domain, specifically claimed  
 in Claim 5(a)".  
 Region 66..91  
 Region "hydrophobic region"  
 101..284  
 /note="mucin-like Ser/Thr-rich region, contains  
 sites for O-linked glycosylation"  
 Domain 285..354  
 /note="EGF-like domain"  
 355..394  
 /note="transmembrane domain"  
 WO9902681-A1.  
 21-JAN-1999.  
 30-JUN-1998; 98WO-US13411.  
 24-JUL-1997; 97US-0899437.  
 09-JUL-1997; 97US-0052019.  
 (GENTH ) GENE TECH INC.  
 Godowski PJ, Mark MR, Zhang D;  
 WPI: 1999-120882/10.  
 N-PSDB; AAX06989.  
 New isolated neuiregulin related ligand-3 - used to develop products  
 for treating nervous system disorders, e.g. stroke, ischemia,  
 infection, malignancy, Alzheimer's disease or Down's syndrome  
 Example 1; Page 78-81; 101pp; English.  
 This is the amino acid sequence of splice variant hNGR32 of human  
 neuiregulin related ligand NRG3, a novel member of the epidermal  
 growth factor (EGF)-like family of protein ligands that binds to  
 the ErbB4 receptor, but not to the ErbB2 or ErbB3 receptor, and  
 which activates ErbB4 receptor tyrosine phosphorylation. The  
 sequence was deduced from the nucleotide sequence of a cDNA clone  
 (see AAX06989) from a foetal brain library. hNGR32 lacks amino  
 acids 529-552 of hNGR3B1 (see AAM97618) but retains the EGF-like  
 domain and is expected to exhibit biological activity. The invention  
 provides human and murine NRG3 polypeptides (see AAM97617), expression  
 vectors, host cells and methods for the recombinant production of  
 NRG3s. The NRG3 polypeptides and polynucleotides and can be used to  
 enhance the survival, proliferation or differentiation of cells  
 having the ErbB4 receptor in vivo and in vitro. They can be used to  
 prevent or treat damage to a nerve or damage to other NRG3-expressing  
 or NRG3-responsive cells, e.g. brain, heart, or kidney cells. In  
 particular, they can be used to treat diseases which involve neural  
 cell growth such as demyelination, or damage or loss of glial cells  
 (e.g. multiple sclerosis). They can be used to treat patients whose  
 nervous system has been damaged by e.g. trauma, surgery, stroke,  
 ischemia, infection, metabolic disease, nutritional deficiency,  
 malignancy, or toxic agents. NRG3 can also be used to treat  
 motor neuron disorders such as amyotrophic lateral sclerosis (Lou  
 Gehrig's disease), Bell's palsy, conditions involving spinal  
 muscular atrophy or paralysis, neurodegenerative disorders such as  
 Alzheimer's disease, Parkinson's disease, epilepsy, multiple  
 sclerosis, Huntington's chorea, Down's syndrome, nerve deafness,  
 and Meniere's disease. They can also be used to treat neuropathies  
 associated with systemic disease including post-polio syndrome,  
 hereditary neuropathies including Charcot-Marie-Tooth disease,  
 Refsum's disease, abetalipoproteinemia, Tangle disease, Krabbe's  
 disease, metachromatic leukodystrophy, Fabry's disease and  
 Dejerine-Sottas syndrome, to treat disease of skeletal muscle of

smooth muscle, such as muscular dystrophy or diseases caused by  
 skeletal or smooth muscle wasting. The products can also be used  
 for detection, diagnosis, for the production of transgenic or  
 knockout animals or for drug screening.  
 Sequence 696 AA;  
 Query Match 100.0%; Score 277; DB 20; Length 696;  
 Best Local Similarity 100.0%; Pred. No. 3e-20;  
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 HFKPCRDRLAYCLNDGECEFYETLTGSHKRCRKEGYQGVRCDFL 47  
 DB 286 hfkpcrdrlayclndgecefyetltgshkrcrckeygyrvrcdfll 332  
 RESULT 6  
 AAM97617  
 ID AAM97617 standard; Protein: 713 AA.  
 AC AAM97617;  
 XX 10-MAY-1999 (first entry)  
 DE Mouse neuiregulin related ligand NRG3.  
 XX  
 KW Neuiregulin related ligand; NRG3; mouse; ErbB4 receptor;  
 KW signal transduction; nervous system disorder; neurodegeneration;  
 KW neuropathy; therapy; diagnosis.  
 XX  
 OS Mus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 1..362  
 FT "note="extracellular domain, specifically claimed  
 in Claim 5(a)".  
 FT Region 66..91  
 FT "note="hydrophobic region"  
 FT 105..286  
 FT "note="mucin-like Ser/Thr-rich region, contains  
 sites for O-linked glycosylation"  
 FT Domain 287..334  
 FT "note="EGF-like domain"  
 FT 363..385  
 FT "note="transmembrane domain"  
 WO9902681-A1.  
 21-JAN-1999.  
 30-JUN-1998; 98WO-US13411.  
 24-JUL-1997; 97US-0899437.  
 09-JUL-1997; 97US-0052019.  
 (GENTH ) GENE TECH INC.  
 Godowski PJ, Mark MR, Zhang D;  
 WPI: 1999-120882/10.  
 N-PSDB; AAX06987.  
 New isolated neuiregulin related ligand-3 - used to develop products  
 for treating nervous system disorders, e.g. stroke, ischemia,  
 infection, malignancy, Alzheimer's disease or Down's syndrome  
 Claim 5(b); Page 59-62; 101pp; English.  
 This is the amino acid sequence of murine neuiregulin related ligand  
 NRG3, a novel member of the epidermal growth factor (EGF)-like  
 family of protein ligands that binds to the ErbB4 receptor, but not  
 to the ErbB2 or ErbB3 receptor, and which activates ErbB4 receptor

tyrosine phosphorylation. The sequence was deduced from the nucleotide sequences of cDNA clones (see AAW05987) from a mouse brain library. The EGF-like domain of NRG3 is distinct from those of NRG1 or NRG2, and NRG3 displays receptor binding characteristics that are distinct from those of other neuroregulins. The invention provides human and murine NRG3 polypeptides (see also AAW97618), expression vectors, host cells and methods for the recombinant production of NRG3. The NRG3 polypeptides and polynucleotides and can be used to enhance the survival, proliferation or differentiation of cells having the ErbB4 receptor in vivo and in vitro. They can be used to prevent or treat damage to a nerve or damage to other NRG3-expressing or NRG3-responsive cells, e.g. brain, heart, or kidney cells. In particular, they can be used to treat diseases which involve neural cell growth such as demyelination, or damage or loss of glial cells (e.g. multiple sclerosis). They can be used to treat patients whose nervous system has been damaged by e.g. trauma, surgery, stroke, ischemia, infection, metabolic disease, nutritional deficiency, malignancy, or toxic agents. NRG3 can also be used to treat motor neuron disorders such as amyotrophic lateral sclerosis (Lou Gehrig's disease), Bell's palsy, conditions involving spinal muscular atrophy or paralysis, neurodegenerative disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, Huntington's chorea, Down's syndrome, nerve deafness, and Meniere's disease. They can also be used to treat neuropathies associated with systemic disease including post-polio syndrome, hereditary neuropathies including Charcot-Marie-Tooth disease, Refsum's disease, abetalipoproteinemia, Tangle disease, Krabbe's disease, metachromatic leukodystrophy, Fabry's disease and Dejerine-Sottas syndrome, to treat disease or diseases caused by skeletal or smooth muscle wasting. The products can also be used for detection, diagnosis, for the production of transgenic or knockout animals or for drug screening.

Sequence 713 AA:

Query Match 100.0%; Score 277; DB 20; Length 713;  
Best Local Similarity 100.0%; Pred. No. 3.1e-20;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 HFKPCRDKLAVCLNDECFVETLTGSHKHCRCKEGYGVRCDOFL 47  
286 hfkpcrdklavclndecfvetltgshkhcrckegygvrddqfl 334

RESULT 7

AAW05452 standard; Protein; 720 AA.

AAW05452;

06-JUL-1999 (first entry)

Human heregulin-like factor sequence.

Human heregulin-like factor; HLF; cell growth regulator; diagnosis;  
neural system disorder; cancer.

Homo sapiens.

MO9857989-A1.

23-DEC-1998.

16-JUN-1998; 98WO-US12403.

17-JUN-1997; 97US-0049942.

(HUMA-) HUMAN GENOME SCI INC.  
(GEOSU) UNIV GEORGETOWN.

Hajazi MM, King CR, Ruben SM, Young P;

WPI: 1999-095327/08.  
New isolated heregulin-like factor - used to develop products for the diagnosis and treatment of disorders involving regulation of cell growth, particularly cancers  
Disclosure; Page 97-99; 118pp; English.  
This sequence is the human heregulin-like factor (HLF) of the invention. The HLF is involved in the regulation of cell growth. Detection of different levels of expression of the HLF gene can be used for the diagnosis of disorders, e.g. in the neural system. In particular, detection of different levels of HLF gene expression in cells or body fluid of an individual can be used for diagnosing cancer. The products can also be used in the treatment of disorders involving abnormal levels of HLF activity.

Sequence 720 AA:

Query Match 100.0%; Score 277; DB 20; Length 720;  
Best Local Similarity 100.0%; Pred. No. 3.1e-20;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 HFKPCRDKLAVCLNDECFVETLTGSHKHCRCKEGYGVRCDOFL 47  
286 hfkpcrdklavclndecfvetltgshkhcrckegygvrddqfl 332

RESULT 8

AAW97618 standard; Protein; 720 AA.

AAW97618;

10-MAY-1999 (first entry)

Human heregulin related ligand NRG3.

Neuregulin related ligand; NRG3; hNRG3B1; human; ErbB4 receptor;  
signal transduction; nervous system disorder; neurodegeneration;  
neuropathy; therapy; diagnosis.

Homo sapiens.

Location/Qualifiers

1..360 "extracellular domain, specifically claimed in Claim 5(a)"

66..91 "hydrophobic region"

101..284 "mucin-like Ser/Thr-rich region, contains sites for O-linked glycosylation"

285..354 "EGF-like domain"

355..394 "transmembrane domain"

WO9902681-A1.

21-JAN-1999.

30-JUN-1998; 98WO-US13411.

24-JUL-1997; 97US-0899437.

09-JUL-1997; 97US-0052019.

(GEENT) GENENTECH INC.  
Godowski PJ, Mark MR, Zhang D;

DR WPI: 1999-120882/10.  
 DR N-PSDB; AAM05182.  
 XX  
 PT New isolated neuregulin related ligand-3 - used to develop products  
 PT for treating nervous system disorders, e.g. stroke, ischemia,  
 PT infection, malignancy, Alzheimer's disease or Down's syndrome  
 XX  
 PS Claim 5(b); Page 66-69; 101pp; English.  
 XX  
 CC This is the amino acid sequence of human neuregulin related ligand  
 CC NR3, a novel member of the epidermal growth factor (EGF)-like  
 CC family of protein ligands that binds to the ErbB4 receptor, but not  
 CC to the ErbB2 or ErbB3 receptor, and which activates ErbB4 receptor  
 CC tyrosine phosphorylation. The sequence was deduced from the  
 CC nucleotide sequence of a cDNA clone (see AAM05182) from a fetal brain  
 CC library. The EGF-like domain of NR3 is distinct from those of NR1  
 CC or NR2, and NR3 displays receptor binding characteristics that are  
 CC distinct from those of other neuregulins. An alternatively spliced  
 CC form of human NR3 is provided in AAM97619. The invention provides  
 CC human and murine NR3 polypeptides (see also AAM97617), expression  
 CC vectors, host cells and methods for the recombinant production of  
 CC NR3s. The NR3 polypeptides and polynucleotides and can be used to  
 CC enhance the survival, proliferation or differentiation of cells  
 CC having the ErbB4 receptor in vivo and in vitro. They can be used to  
 CC prevent or treat damage to a nerve or damage to other NR3-expressing  
 CC or NR3-responsive cells, e.g. brain, heart, or kidney cells. In  
 CC particular, they can be used to treat diseases which involve neural  
 CC cell growth such as demyelination, or damage or loss of glial cells  
 CC (e.g. multiple sclerosis). They can be used to treat patients whose  
 CC nervous system has been damaged by e.g. trauma, surgery, stroke,  
 CC ischemia, infection, metabolic disease, nutritional deficiency,  
 CC malignancy, or toxic agents. NR3 can also be used to treat  
 CC motor neuron disorders such as amyotrophic lateral sclerosis (Lou  
 CC Gehrig's disease), Bell's palsy, conditions involving spinal  
 CC muscular atrophy or paralysis, neurodegenerative disorders such as  
 CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple  
 CC sclerosis, Huntington's chorea, Down's syndrome, nerve deafness,  
 CC and Meniere's disease. They can also be used to treat neuropathies  
 CC associated with systemic disease including post-polio syndrome,  
 CC hereditary neuropathies including Charcot-Marie-Tooth disease,  
 CC Refsum's disease, abetalipoproteinemia, Tangier disease, Krabbe's  
 CC disease, metachromatic leukodystrophy, Fabry's disease and  
 CC Dejerine-Sottas syndrome, to treat disease of skeletal muscle of  
 CC smooth muscle, such as muscular dystrophy or diseases caused by  
 CC skeletal or smooth muscle wasting. The products can also be used  
 CC for detection, diagnosis, for the production of transgenic or  
 CC knockout animals or for drug screening.  
 CC  
 CC Sequence 720 AA;  
 SO

Query Match 100.0%; Score 277; DB 20; Length 720;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-20;  
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 HFKPRDLDVAVCLNDGECFVETLTGSHKH-CRCKEGYGVRCDOFL 47  
 Db 286 hfkprdkldvavclndgectvletltgshkncrckegygvrctdfl 332

RESULT 9  
 ID AAM05182 standard; peptide: 52 AA.  
 AC AAM05182;  
 XX  
 DT 04-JUN-1997 (first entry)  
 XX  
 DE Neu differentiation factor/hergulin-alpha/beta form EGF-like domain.  
 XX  
 KW NDF; neu differentiation factor; heregulin; epidermal growth factor;  
 KW EGF; colon epithelial cell proliferation; Schwann cell; nerve;  
 KW damage; colitis; ulcer.  
 DR

XX  
 OS Synthetic.  
 XX  
 PN WO963159-A1.  
 XX  
 PD 10-OCT-1996.  
 XX  
 PF 27-MAR-1996; 96WO-US04262.  
 XX  
 PR 06-APR-1995; 95US-0417640.  
 XX  
 PA (AMGE-) AMGEN INC.  
 XX  
 PI Cranahan JF, Hara S, Lu HS, Mayer JP, Yoshinaga SK;  
 XX  
 DR WPI: 1996-465022/46.  
 XX  
 PT Peptide(s) derived from neu differentiation factor/hergulin  
 PT proteins - specifically from epidermal growth factor-like domain,  
 PT stimulate proliferation of colon epithelial cells and Schwann cells  
 XX  
 PS Claim 1; Page 27; 37pp; English.  
 CC  
 CC The peptides AAM05182-W05185 are based on neu differentiation factor  
 CC (NDF)/hergulin alpha and beta form EGF-like domains in various  
 CC combinations. The peptides maintain the survival and proliferation of  
 CC Schwann cells and cause proliferation, growth and differentiation of  
 CC colon epithelial cells. Accordingly, they are useful to treat (in vitro  
 CC or in vivo) a disease or disorder of the colon (e.g. colitis or an  
 CC ulcer) or of the nervous system (e.g. nerve damage caused by trauma).  
 CC  
 CC Sequence 52 AA;  
 SO

Query Match 42.1%; Score 116.5; DB 17; Length 52;  
 Best Local Similarity 33.3%; Pred. No. 4.7e-05;  
 Matches 16; Conservative 15; Mismatches 16; Indels 1; Gaps 1;

Oy 1 HFKPRDLDVAVCLNDGECFVETLTGSHKH-CRCKEGYGVRCDOFL 47  
 Db 2 hlvkaeketkctvngccfmwkdlnspyrlyckqpfgtgarcqnyv 49

RESULT 10  
 ID AAB12602 standard; peptide: 52 AA.  
 AC AAB12602;  
 XX  
 DT 09-NOV-2000 (first entry)  
 XX  
 DE Human NDF EGF-like domain derived peptide SEQ ID NO:1.  
 XX  
 KW Human; sensory epithelial cell; growth; stimulant; inner ear; EGF;  
 KW epithelial growth factor; NDF; heregulin; monoclonal antibody;  
 KW adult rat utricular epithelium.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6080845-A.  
 XX  
 PD 27-JUN-2000.  
 XX  
 PE 28-JAN-1999; 99US-0238182.  
 XX  
 PR 05-AUG-1998; 98US-0129549.  
 XX  
 PA (AMGE-) AMGEN INC.  
 XX  
 PI Cranahan JF;  
 XX  
 DR WPI: 2000-451228/39.  
 DR N-PSDB; AAM05182.  
 DR



```
DE      EGF12.
XX
KW      Glial growth factor; GGF, heregulin; mitogenesis;
KW      Schwann cell; tumour; central nervous system; erbB2 receptor;
KW      antiproliferative; epidermal growth factor; EGF.
XX      WO9403644-A.
XX
XX      PD
XX      17-FEB-1994.
XX
PF      10-AUG-1993;    93WO-US07491.
XX
PR      10-AUG-1992;    92US-0927337.
PR      25-SEP-1992;    92US-0951747.
PR      01-DEC-1992;    92US-0984085.
PR      29-JAN-1993;    93US-0011396.
XX
PA      (CAMP-) CAMBRIDGE NEUROSCIENCE INC.
XX
PI      Gwynne DI, Marchionni M, McBurney RN;
XX      WPI. 1994-065732/08.
DR      N-PDB; AAQ58324.
XX
PT      Glial growth factor DNA encoding numerous polypeptide factors
PT      used for inhibiting cell proliferation - for treating carcinoma
PT      and nervous disorders
XX
PS      Disclosure: Fig 40; 17pp; English.
XX
XX      The GGF coding segments include regions with EGF-like homology.
CC      These EGF-like domains can be required for the activation of
CC      mitogenesis in the binding reaction between GGF ligands contg.
CC      such domains and the erbB2 receptor. Pref. antiproliferative
CC      factors are those which lack these EGF-like domains.
XX
SO      Sequence     63 AA;

Query Match          41.0%; Score 113.5; DB 15; Length 63;
Best Local Similarity 34.8%; Pred. No.0.00011;
Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1.

OY      1 HFKPGRKDLAVCLNDGECVEIVETLTGSCHK-CRCKEGYGVRCDQ 45
        | : : : : : ||||| : : : : : || :
Dd      2 hlvkaeketfcvngsecfmwkdlnpsrylckcqpftgarcte 47

RESULT 14
AAR55659
ID      AAR55659 standard; Protein; 63 AA.
XX
AC      AAR55659;
XX
XX      28-JUL-1994 (first entry)
XX
DE      EGFL2.
XX
KW      Glial growth factor; GGF, heregulin; mitogenesis;
KW      Schwann cell; tumour; central nervous system;
KW      epidermal growth factor; EGF.
XX
PN      WO9400140-A.
XX
PD      06-JAN-1994..
XX
PE      29-JUN-1993;    93WO-US06228.
XX
PR      30-JUN-1992;    92US-0907138.
PR      03-SEP-1992;    92US-0940389.
PR      23-OCT-1992;    92US-0965173.
PR      24-MAR-1993;    93US-0036555.
XX
```

PA	(CAMP-) CAMBRIDGE NEUROSCIENCE.
PA	(LUDW-) LUDWIG INST CANCER RES.
XX	
XX	Chen MS, Goodearl A, Hiles I, Marchionni M, Minghetti L;
PI	Stroobant P, Waterfield M;
XX	
DR	WPI; 1994-025882/03.
DR	N-PSDB; AAQ62843.
XX	
PT	Glia mitogenic polypeptide factors - useful for stimulating
XX	glial cell mitogenesis and treating glial cell tumours
XX	
PS	Claim 53; Fig 39; 178pp; English.
CC	EGFL1, EGFL2, EGFL3, EGFL4, EGFL5 and EGFL6 are used
CC	for the stimulation of glial cell mitogenesis in vivo
CC	and in vitro.
XX	
SQ	Sequence 63 AA.
OY	
Dd	Query Match 41.0%; Score 113.5; DB 15; Length 63; Best Local Similarity 34.8%; Pred. No. 0.00011; Matches 16; Conservative 14; Mismatches 15; Indels 1; Gaps 1;
1	HFKPCRDLAYCLNDGECVETITGSHKH-CRCKEGYQGVRCDQ 45 1 : 2 hlvkaeketfcvngsecfmwkdlpsrylylcqcpftgarcte 47
RESULT 15	
AAR67250	
ID	AAR67250 standard; Protein; 63 AA.
AC	
AA67250;	
DT	15-AUG-1995 (first entry)
DE	
XX	Human epidermal like growth factor 2 (EGFL2).
XX	
KM	Epidermal like growth factor 2; mammalian muscle cell treatment;
KW	skeletal; cardiac; smooth; acetylcholine receptor deficiency;
KM	EGFL2.
XX	
OS	Homo sapiens.
XX	
PN	MO9426298-A.
PD	
XX	24-NOV-1994.
PF	
XX	06-MAY-1994; 94WO-US05083.
PR	
XX	06-MAY-1993; 93US-0059022.
PR	08-MAR-1994; 94US-0209204.
XX	
PA	(CAMP-) CAMBRIDGE NEUROSCIENCE.
XX	
PI	Gwynne DI, Marchionni M, Sklar R;
XX	
DR	WPI; 1995-006353/01.
DR	N-PSDB; AAQ74915.
PT	Treating mammalian muscle diseases and disorders - by admin. of
PT	GGP2 and other specified polypeptide(s) which bind the p185erbB2
PT	receptor.
XX	
PS	Claim 34; Pages 148-149; 241pp; English.
XX	
AAQ74915	encodes AAR67250 human epidermal like growth factor 2 (EGFL2).
CC	The glial cell mitogenic activity of EGFL2 can be used to treat a
CC	variety of mammalian skeletal, cardiac and smooth muscle diseases,
CC	including acetylcholine receptor deficiency.
XX	



